

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAACAGAAAACCTGTTAGAAATGTGGTGGT
TTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTT
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGGCAGTTTTATGCATTG
CTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAAA
TTAAACAAGGCTGGCCTTGTAAGTGGAACTGAGTTGTTTAGGACTTCTATTGTGGCAAACCTT
CCAGAAAACAACCCTTTTTGTGCACATGTAAGTGGAGCTGTGCTTACCTTTGGTATGGGCTCAT
TATATATGTTTGTTCAGACCATCCTTTCCTACCAAATGCAGCCCAAAATCCATGGCAAACAAGTC
TTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGGAGTAAGTGCACCTTAGCATGCTGACTTGCTC
ATCAGTTTTGCACAGTGGCAATTTTGGGACTGATTTAGAACAGAACTCCATTGGAACCCCGAGG
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAATGGTCTATGTCATTTTCCTTCTTT
GGTTTTTTCCTGACTTACATTCGTGATTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACA
TGGATTAACCTCTATGACACTGCACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCA
GAGATATTTGATGAAAGGATAAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGG
TTCACAGAAGTTGCTTATTCTTCTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCAT
CAAGAAGACTATTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIIFSITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMOPKIHGKQVFWIRLLLVIWCGVSALMLTCSVVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFFLTYYIRDFQKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 4

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV
 LNSSSRYPFHWMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTfMYFFWKLGDp
 FPILSPKHGILSIEQLISRVGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDILALERRLLQ
 TMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQQEVDAL EELSRQ
 LFLETADLYATKERIEYSKTFKGKYFNLGYFFSIYCVWKIFMATINIVFDRVGKTDpVTRGIEI
 TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLLAQIMGMY
 FVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
 425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCTCGGTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGTTCC
 AGAACTCTCCATCCGGACTAGTTATGAGCATCTGCCTCTCATATCACCAGTGGCCATCTGAGGT
 GTTTCCTTGGCTCTGAAGGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTGCTTCTCACT
 TCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCTTTCCATCCAGGT
 GTCATGCAGAATTATGGGGATCACCTTTGTGAGCAAAAAGGCGAACCAGCAGCTGAATTTACAG
 AAGCTAAGGAGGCCTGTAGGCTGCTGGGACTAAGTTTGGCCGCAAGGACCAAGTTGAAACAGCC
 TTGAAAGCTAGCTTTGAACTTGCAGCTATGGCTGGGTTGGAGATGGATTCGTGGTCATCTCTAG
 GATTAGCCCAAAACCCCAAGTGTGGGAAAAATGGGGTGGGTGTCCTGATTTGGAAGGTTCCAGTGA
 GCCGACAGTTTGCAGCCTATTGTTACAACCTCATCTGATACTTGGACTAATCGTGCATTCCAGAA
 ATTATCACCACCAAGATCCCATATTTCAACACTCAAACCTGCAACACAAACACAGAATTTATTGT
 CAGTGACAGTACCTACTCGGTGGCATCCCCCTTACTCTACAATACCTGCCCCCTACTACTCTCTC
 CTGCTCCAGCTTCCACTTCTATTCCACGGAGAAAAAATTGATTTGTGTACAGAAAGTTTTATG
 GAACTAGCACCATGTCTACAGAACTGAACCATTGTTGAAAATAAAGCAGCATTCAGAATGA
 AGCTGCTGGGTTTGGAGGTGTCCCCACGGCTCTGCTAGTGCTTGTCTCTCTCTCTCTCTCTCTG
 CAGCTGGTCTTGGATTTTGTCTATGTCAAAAGGTATGTGAAGGCCTTCCCTTTTACAAACAAGAAT
 CAGCAGAAGGAAATGATCGAAACCAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCCATA
 TGAGGAATCAAAGAAACTGATAAAACCCAGAAAGTCCAAGAGTCCAAGCAAAACTACCGTGC
 GATGCCCTGGAAGCTGAAGTTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTTTCTTT
 CATGCTCCTTACCCTGCCCCAGCTGGGGAAATCAAAAGGGCCAAAGAACCAGAAAGAAAGTCCA
 CCCTTGGTTCCTAACTGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAGAGAAT
 GCCCTTCTCCTTATTGTAACCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCCACGGCCT
 TTCTAGCCTGGCTATGTCTTAATAATATCCCCTGGGAGAAAGGAGTTTTGCAAGTGCAGGAC
 CTAACACATCTCATCAGTATCCAGTGGTAAAAAGGCCTCCTGGCTGTCTGAGGCTAGGTGGGTG
 AAAGCCAAGGAGTCACTGAGACCAAGGCTTTCTCTACTGATTCCGAGCTCAGACCCCTTTCTTCA
 GCTCTGAAAGAGAAACACGTATCCACCTGACATGTCTTCTGAGCCCGGTAAAGAGCAAAAGAAAT
 GGCAGAAAAGTTTAGCCCTGAAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTCTGTAAA
 GCTAAAATAAAGAAATAGAACAAGGCTGAGGATACGACAGTACACTGTGAGCAGGGACTGTAAC
 ACAGACAGGGTCAAAGTGTTTTCTCTGAACACATTGAGTTGGAATCACTGTTTAGAACACACACA
 CTTACTTTTTCTGGTCTCTACCCTGCTGATATTTCTCTAGGAAATATACTTTTACAAGTAACA
 AAAATAAAACTCTTATAAATTTCTATTTTATCTGAGTTACAGAAATGATTACTAAGGAAGATT
 ACTCAGTAATTTGTTTAAAAAGTAATAAAATTCACAAACATTTGCTGAATAGCTACTATATGTC
 AAGTGCTGTGCAAGGTATTACACTCTGTAATTGAATATTATTCCTCAAAAAATTGCACATAGTAG
 AACGCTATCTGGGAAGCTATTTTTTTCAGTTTGTATTTCTAGCTTATCTACTTCCAACTAAT
 TTTTATTTTGTGCTGAGACTAATCTTATTCATTTTCTCTAATATGGCAACCATTATAACCTTAAT
 TATTATTAACATACCTAAGAAGTACATTGTTACCTCTATATACCAAGCACATTTTAAAGTGCC
 ATTAACAAATGTATCACTAGCCCTCCTTTTTCCAACAAGAGGACTGAGAGATGCAGAAATATT
 TGTGACAAAAAATTAAAGCATTTAGAAACTT

FIGURE 6

MARCFSLVLLLSIWTTTRLLVQGSRLAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG
LSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYN
SSDTWTNSCIPEIIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPPAPASTSIPR
RKKLICVTEVFMETSTMSTETEPFVENKAAFKNAAAGFGGVPTALLVLALLFFGAAAGLGFCYVK
RYVKAFPFTNKNQKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCGCGCTCCCGCACCCGCGGGCCCGCCACCGCGCGCTCCCGCATCTGCACCCGCGAGCCCGGC
GGCCTCCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCGCAGCGCAACTCGGTCCAGTCCGGGCGG
CGGCTGCGGGCGCAGAGCGGAGATGAGCAGCGGCTTGGGGCCACCCTGCTGTGCCTGCTGCTGGCGG
CGGCGGTCCCCACGGCCCCCGCGCCCGCTCCGACGGCGACCTCGGCTCCAGTCAAGCCCGGCCCG
GCTCTCAGCTACCCGCGAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGAGGTTGAGGAAGTAT
GGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCGAGAAGCTGCTGCTA
AAGCATCATCAGAAGTGAACCTGGCAAACCTTACCTCCAGCTATCACAATGAGACCAACACAGAC
ACGAAGGTTGGAAATAATACCATCCATGTGCACCGAGAAATTCACAAGATAACCAACAACAGAC
TGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGACGAAGAAGGCAGAAGGAGCC
ACGAGTGCATCATCGACGAGGACTGTGGGCCCAGCATGTACTGCCAGTTTGGCAGCTTCCAGTAC
ACCTGCCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAGTGCTGTGGAGACCA
GCTGTGTGTCTGGGGTCACTGCACCAAAATGGCCACCAGGGGCGAGCAATGGGACCATCTGTGACA
ACCAGAGGGACTGCCAGCCGGGGCTGTGCTGTGCCTTCCAGAGAGGCTGCTGTTCCCTGTGTGC
ACACCCCTGCCCAGTAAAGTTTCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTGTGTCAGCT
CTGGGAGCTAGAGCCTGATGGAGCCTTGGACCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGC
CCCACAGCCACAGCCTGGTGTATGTGTGCAAGCCGACCTTCGTGGGAGCCGTGACCAAGATGGG
GAGATCCTGCTGCCAGAGAGGTCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCG
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTG
CCGCCGCTGCACTGCTGGGAGGGGAAGAGATTAGATCTGGACCAGGCTGTGGGTAGATGTGCAA
TAGAAATAGCTAATTTATTTCCCGAGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCCTA
CATCTTCTTCCCAGTAAGTTTCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTGTGTCAGCT
CCCCCAGGCTGTTCTCCAGGCTTACAGTCTGGTGTCTGGGAGAGTCAGGCAGGGTTAACTGCA
GGAGCAGTTTGGCACCCTGTCCAGATTATTGGCTGCTTTGCCCTTACCAGTTGGCAGACAGCCG
TTTGTCTACATGGCTTTGATAATTGTTTGGAGGGAGGAGATGGAAACAATGTGGAGTCTCCCTC
TGATTGGTTTTGGGGAATGTGGAGAAGAGTGCCCTGCTTTGCAAAACATCAACCTGGCAAAAATG
CAACAAATGAATTTTCCACGCAGTTCTTCCATGGGCATAGGTAAGCTGTGCCTTACAGTGTGTC
AGATGAAATGTTCTGTTACCCCTGCATTACATGTGTTTATTCATCCAGCAGTGTGCTCAGCTCC
TACCTCTGTGCCAGGGCAGCATTTTCATATCCAAGATCAATTCCCTCTCTCAGCACAGCCTGGGG
AGGGGGTCAATTGTTCTCCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTTGCC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCACTCTGGTTGTGACTCTAAGCTCAGTGCTCT
CTCCACTACCCACACCGCCTTGGTGCCACCAAAAGTGCTCCCCAAAAGGAAGGAGAATGGGAT
TTTTCTTGAGGCATGCACATCTGGAATTAAGGTCAAACATAATCTCACATCCCTCTAAAAGTAAA
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGCGAGCGTCCTTCTAATGAAGACAATGAT
ATTGACACTGTCCCTCTTGGCAGTTGCATTAGTAACCTTGAAAGGTATATGACTGAGCGTAGCA
TACAGGTTAACTGCAGAAACAGTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGC
AAAATCACTTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC
TGTGTGAAACATGGTTGTAATATGCGACTGCGAACACTGAACCTTACGCCACTCCACAAATGATG
TTTTCAGGTGTGATGACTGTTGCCACCATGTATTCATCCAGAGTCTTAAAGTTAAAGTTGCA
CATGATTGTATAAGCATGCTTTCTTTGAGTTTAAATATGTATAAACAATAAGTTGCATTTAGAA
ATCAAGCATAAATCACTTCAACTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCILLAAVPTAPAPAPTATSAPVKPGPALSYQEEATLNEMFREVEELMEDTQHKL
RSAVEEMEAEAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTGQMFSE
TVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQRLCTRDSECCGDQLCVWGHC
TKMATRGSNGTICDNQRDCQPGLCCAFQRGLLFPVCTPLPVEGELCHDPASRLDLITWELEPDG
ALDRPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQELEDLE
RSLTEEMALGEPAAAAAALLGGEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCCAC
 GGCCACCTTGTGAACCTCGTGCCAGGGCTGATGTGCGTCTCCAGGGCTACTCATCAAAG
 GCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCTGGGGCTCTTCTGGACCTT
 AACTGGGTACTGGCCCTGGGCCAATGCGTCTCGCTGGAGCCTTGCCTCCTTCTACTGGGCCTT
 CCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCCGCACACTCCGTTACC
 ACACTGGGTCATTGGCATTGGAGCCCTCATCTGACCCTTGTGCAGATAGCCCGGTCATCTTG
 GAGTATATTGACCACAAGCTCAGAGGAGTGACAGAACCTGTAGCCCGCTGCATCATGTGCTGTTT
 CAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCCTAAACCGCAATGCATACATCATGA
 TCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAAAATGCGTTCATGCTACTCATGCGAAAC
 ATTGTCAGGGTGGTCGTCTCGGACAAAGTCACAGACCTGCTGCTGTTCTTTGGGAAGCTGCTGGT
 GGTGCGAGGCGTGGGGTCTGTCTTCTTTTTTCTCCGGTCGCATCCCGGGGCTGGGTAAAG
 ACTTTAAGAGCCCCACCTCAACTATTACTGGCTGCCCATCATGACCTCCATCCTGGGGGCTAT
 GTCATCGCCAGCGGCTTCTTCAGCGTTTTCGGCATGTGTGTGGACACGCTCTTCTCTGCTTCTT
 GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAA
 AGATTCTGGGCAAGAAGAACGAGGCGCCCCGACAACAAGAAGAGGAAGAAGTGACAGCTCCGG
 CCTGATCCAGGACTGCACCCACCCCAACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGT
 CTCCATTTTGTGGTAAAAAAGGTTTTAGGCCAGGCGCGTGGCTCACGCCTGTAATCCAACACT
 TTGAGAGGCTGAGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTG
 AAACCTCCGTCTCTATTAAAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCA
 GCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGA
 GATCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAAACAAA
 AAGATTTTATTAAAGATATTTTGTTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLEFWTL
 NWVLALGQCVLGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
 EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAFMLLMRN
 IVRVVLDKVTDLLLFFGKLLVVGGVVLSFFFFSGRIPGLGKDFKSPHLNYYWLPIMTSILGAY
 VIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYMSKSLKILGKKNEAPPDNKKRKK

Important features:**Transmembrane domains:**

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 12

MGACLGACSLSCASCLCGSAPCILSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVESQL
YKLPWVCEEAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFTLLMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHSWNQRWLGKAE
ECDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFVCVVSIAAVLPKV
QDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI
VGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQVAACEGRAFDNEQDGVITYSY
SFFHFCLVLASLHVMMTLTNWYKPGETRMISTWTAVWVKICASWAGLLLYLWTLVAPLLLRNRD
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGGCCGCCAGGAACCAACCCGTTAAGGTGTCTTCTCTTTAGGGATGGTGA
 GGTGGAAGAAAGACTCCTGTAACCTCCTCCAGGATGAACCACTGCCAGAAGACATGGAGAACG
 CTCTACCGGGGAGCCAGAGCTCCCATGCTTCTCTGCGCAATATCCATTCCATCAACCCACACAA
 CTCATGGCCAGGATTGAGTCCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
 TTTCTGTTTGTGTTGTACCTTTGACCTCTTATTCGTAACATTACTGTGGATAATAGAGTTAAATG
 TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTTCATAT
 TTTGATATATTTCTTCTGGCAGTTTTTCGATTTAAAGTGTTAATACTTGCATATGCTGTGTGCAG
 ACTGCGCCATTGGTGGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTTTACTAGCAAAAG
 TGATCCTTTTCGAAGCTTTTCTCTCAAGGGGCTTTTGGCTATGTGCTGCCCATCATTTCATTATC
 CTTGCCTGGATTGAGACGTGGTTCCTGGATTTCAAAGTGTTACCTCAAGAAGCAGAAGAAGAAAA
 CAGACTCCTGATAGTTCAGGATGCTTCAGAGAGGGCAGCACTTATACCTGGTGGTCTTTCTGATG
 GTCAGTTTTATTCCCCTCCTGAATCCGAAGCAGGATCTGAAGAAGCTGAAGAAAAACAGGACAGT
 GAGAAACCACTTTTAGAACTATGAGTACTACTTTTTGTTAAATGTGAAAAACCCTCACAGAAAGTC
 ATCGAGGCAAAAAGAGGCAGGCAGTGGAGTCTCCCTGTGACAGTAAAGTTGAAATGGTGACGTC
 CACTGCTGGCTTTATTGAACAGCTAATAAAGATTTATTTATTGTAATACCTCACAAACGTTGTAC
 CATATCCATGCACATTTAGTTGCCTGCCTGTGGCTGGTAAGGTAATGTCATGATTCATCCTCTCT
 TCAGTGAGACTGAGCCTGATGTGTTAACAAATAGGTGAAGAAAGTCTGTGCTGTATTCTTAATC
 AAAAGACTTAATATATTGAAGTAACACTTTTTTAGTAAGCAAGATACCTTTTTATTTCATTCAC
 AGAATGGAATTTTTTTGTTTCATGTCTCAGATTTATTTTGATTTCTTTTTTAACACTCTACATT
 TCCCTTGTTTTTTAACTCATGCACATGTGCTCTTTGTACAGTTTTAAAAAGTGAATAAAATCTG
 ACATGTCAATGTGGCTAGTTTTATTTTCTTGTTTTGCATTATGTGTATGGCCTGAAGTGTGGGA
 CTTGCAAAAGGGGAAGAAAGGAATTGCGAATACATGTAAAATGTCACCAGACATTTGTATTATTT
 TTATCATGAAATCATGTTTTCTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTTTGAATGC
 ACAAATGACTTAAACCATTATATCATGTTTCCTTTGCGTTCAGCCAATTTCAATTTAAATGAA
 CTAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLF
VTLLWIIELNVNGGIENTLEKEVMQYDYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT
AVTSAFLLAKVILSKLFSQGAFGYVLPPIISFILAWIETWFLDFKVLQPQAEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEEAEKQDSEKPLEL

Important features of the protein:**Signal peptide:**

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 17

GCAGCGGCGAGGCGGCGGTGGTGGCTGAGTCCGTGGTGGCAGAGGCGAAGGCGACAGCTCATGCG
GGTCCGATAGGGCTGACGCTGCTGCTGTGTGCGGTGCTGCTGAGCTTGGCCTCGGCGTCCTCGG
ATGAAGAAGGCAGCCAGGATGAATCCTTAGATTCCAAGACTACTTTGACATCAGATGAGTCAGTA
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTTCTTGATTGAGAAGAATCTGA
ATTAGAATCCTCTATTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGAAAGTGTCACAG
AAGATATCAGCTTTCTAGAGTCTCCAAATCCAGAAAACAAGGACTATGAAGAGCCAAAGAAAGTA
CGGAAACCAGCTTTGACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTTTCT
TTTCCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGGCAGACTGTGGTGTG
CTACAACCTATGACTACAAAGCAGATGAAAAGTGGGGCTTTTGTGAACTGAAGAAGAGGCTGCT
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAACCTGGAATGAAAATCCTTAATGGAAG
CAATAAGAAAAGCCAAAAAGAGAAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA
CCAAAGCCCTGGAGAGAGTGTCTATGCTCTTTTATTTGGTGATTACTTGCCACAGAATATCCAG
GCAGCGAGAGAGATGTTTGAGAAGCTGACTGAGGAAGGCTCTCCCAAGGACAGACTGCTCTTGG
CTTTCTGTATGCCTCTGGACTTGGTGTTAATTCAAGTCAGGCAAAGGCTCTTGATATTATACAT
TTGGAGCTCTTGGGGCAATCTAATAGCCCACATGGTTTTTGGTAAGTAGACTTTAGTGGAAGGCT
AATAATATTAACATCAGAAGAATTTGTGGTTTATAGCGGCCACAACCTTTTTCAGCTTTTCATGATC
CAGATTTGCTTGTTATTAAGACCAAATATTCAGTTGAACTTCCTTCAAATCTTGTTAATGGATAT
AACACATGGAATCTACATGTAAATGAAAGTTGGTGGAGTCCACAATTTTCTTTAAATGATTAG
TTTGGCTGATTGCCCCATAAAAGAGAGATCTGATAAATGGCTCTTTTAAATTTTCTCTGAGTTG
GAATTGTCAGAATCATTTTTTACATTAGATTATCATAATTTTAAAAATTTTCTTTAGTTTTTCA
AAATTTTGTAATGGTGGCTATAGAAAAACAACATGAAATATTATACAATATTTGCAACAATGC
CCTAAGAATTGTTAAATTCATGGAGTTATTTGTGCAGAATGACTCCAGAGAGCTCTACTTTCTG
TTTTTTACTTTTCATGATTGGCTGTCTTCCATTTATTCTGGTCATTTATTGCTAGTGACACTGT
GCCTGCTTCCAGTAGTCTCATTTTCCCTATTTTGCTAATTTGTTACTTTTTCTTTGCTAATTTGG
AAGATTAACTCATTTTTTAATAAAATTATGTCTAAGATTAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 18

MRVRIGLTLLLCVLLSLASASSDEEGSQDES LDSKTTLTSDSVKDHTTAGRVVAGQIFLDSESESEL
 ESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEP KVRKPALTAIEGTAHGEPCHF PFLFLDK
 EYDECTSDGREDGRLWCATTYDYKADEK WGFCEEEEEAAKRRMQEAEMMYQTGMKILNGSNKKSQKR
 EAYRYLQKAASMNHTKALERVSYALLFGDYLPQNIQAAREMF EKLTEEGSPKGQTALGFLYASGLGVN
 SSQAKALVYYTTFGALGGNLI AHMVLVSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCAGATTTTAAGCCCATCTGCAGTGGAATTCATGAAC TAGCAAGAGGACACCATCTTCTT
GTATTATACAAGAAAGGAGTGACCTATCACACACAGGGGGAAAAATGCTCTTTTGGGTGCTAGG
CCTCCTAATCCTCTGTGGTTTTCTGTGGACTCGTAAAGGAAAAC TAAAGATTGAAGACATCACTG
ATAAGTACATTTTTATCACTGGATGTGACTCGGGCTTTGGAACTTGGCAGCCAGAAC TTTTGAT
AAAAAGGGATTTTCATGTAATCGCTGCCTGTCTGACTGAATCAGGATCAACAGCTTTAAAGGCAGA
AACCTCAGAGAGACTTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGAATGTCAAGAGGACTG
CCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGGTCTCTGGGGTCTGATCAATAATGCTGGTGTT
CCCGGCGTGCTGGCTCCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA
CCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTCCTTTGGTCAAGAAAGCTCAAGGGAGAG
TTATTAATGTCTCCAGTGTTGGAGGTCGCCTTGCAATCGTTGGAGGGGGCTATACTCCATCCAAA
TATGCAGTGGAAGGTTTCAATGACAGCTTAAGACGGGACATGAAAGCTTTTGGTGTGCACGTCTC
ATGCATTGAACCAGGATTGTTCAAACAAACTTGGCAGATCCAGTAAAGGTAATTGAAAAAAAC
TCGCCATTTGGGAGCAGCTGTCTCCAGACATCAAACAACAATATGGAGAAGGTTACATTGAAAAA
AGTCTAGACAAACTGAAAGGCAATAAATCCTATGTGAACATGGACCTCTCTCCGGTGGTAGAGTG
CATGGACCACGCTCTAACAAGTCTCTTCCCTAAGACTCATTATGCCGCTGGAAGATGCCAAAA
TTTTCTGGATACCTCTGTCTCACATGCCAGCAGCTTTGCAAGACTTTTTATTGTTGAAACAGAAA
GCAGAGCTGGCTAATCCCAAGGCAGTGTGACTCAGCTAACCACAAATGTCTCTCCAGGCTATGA
AATTGGCCGATTTCAAGAACACATCTCCTTTTCAACCCCATTCCTTATCTGCTCCAACCTGGACT
CATTTAGATCGTGCTTATTTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTGGTATCCAGGGT
CCCTGCTCAAGTTTTCTTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCCTGCCCT
GTATTTAGGCTTTGCCTGCTTGGTGTGATGTAAGGGAAATTGAAAGACTTGCCCATTCAAAATGA
TCTTTACCGTGGCCTGCCCCATGCTTATGGTCCCCAGCATTTACAGTAACTTGTAATGTTAAGT
ATCATCTCTTATCTAAATATTAAAAGATAAGTCAACCCAAAAA AAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRKGKLIKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLTESG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNVQGEKGLWGLINNAGVPGVLAPTDWLTLEDY
REPIEVNLFGLISVTNLMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMK
AFGVHVSCIEPGLFKTNLADPVKVIEKKLAIWEQLSPDIKQQYGEGYIEKSLDKLKGNKSYVNMD
LSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQKAELANPKAV

Important features of the protein:**Signal peptide:**

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGCGGCGGTAGCAATGGAGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCGGCG
 CACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAAGTAAAA
 GGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTATACAATTGA
 CATTTCAGAAATATATTCCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAGGCGAAGTAAATG
 AGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGTTGGTACAAATTCCGT
 CGTCATTTCAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAAAACCTTGCAGGAGCATT
 TTCAAACCAAGACCTTGTTTTCTGCTATTAACACCAAGTATAATAACAGAAAGCTGCTCTACTC
 ATCGACTGGAACATTCTTATATAAACCTCAAAAAGGACTTTTTTCACAGGGTACCTTTAGTGGTT
 GCCAATCTGGGCATGCTGAACAACGGGTATATAAACTGTATCAGGTTCTGTATGTCCACTGG
 TTTTAGCCGAGCAGTACAAAACACAGCTCTAAATTTTTGAAGAAGATGGATCCTTAAAGGAGG
 TACATAAGATAAATGAAATGTATGCTTCATTACAAGAGGAATTAAAGAGTATATGCAAAAAGTG
 GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAAACAGATTAAAACGAGAAATTGA
 GAAAAGGAGAGGAGCACAGATTCAGGCAGCAAGAGAGAAGAACATCCAAAAGACCCTCAGGAGA
 ACATTTTTCTTTGTCAGGCATTACGGACCTTTTTTCCAAATCTGAATTTCTTCATTCATGTGTT
 ATGTCTTTAAAAAATAGACATGTTTCTAAAAGTAGCTGTAACACACCACCATCTCGATGTAGT
 AGACAATCTGACCTTAATGGTAGAACACACTGACATTCTGAAGCTAGTCCAGCTAGTACACCAC
 AAATCATTAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTTGTTA
 GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAT
 GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTGGTGAATATTACGGTCTC
 CTACATTTTGATCTTTTTAACCTTACAAGGAGATTTTTTTATTTGGCTGATGGGTAAAGCCAAAC
 ATTTCTATTGTTTTTACTATGTTGAGCTACTTGCAGTAAGTTCATTTGTTTTTACTATGTTTACC
 TGTTTGAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAACATCA
 GATGCTTTTATTTCCAAACCTTTTTTTCACCTTTCACTAAGTTGTTGAGGGGAAGGCTTACACAG
 ACACATTCTTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCAGCACT
 TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGGCAACGTATT
 GAGACCATGTCTATTAAAAAATAAAATGGAAAAGCAAGAATAGCCTTATTTTCAAATATGAAA
 GAAATTTATATGAAAATTTATCTGAGTCATTAATTTCTCCTTAAGTGATACTTTTTTAGAAGTA
 CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAAATTTGCAAAACATCATCT
 AAAATTTAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFOHLNTDSDEGFLGGEVKGEAKNSITDSQMDDVEVVYTIDIQYI
PCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEHFSNQDL
VFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSRAV
QTHSSKFFEEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVNRLKREIEKRRGA
QIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCYNHHLVDVVDNLTL
MVEHTDIPEASPASTPQIIKHKALDLDLRWQFKRSRLDQDKRSKANTGSSNQDKASKMSSPET
DEEIEKMKGFGEYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGGCGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGCCCAA
GCAGCGCGCAGCGAACGCCCGCGCCGCCACACCCCTCTGCGGTCCCGCGGCGCCTGCCACCCCTCCCTCCTTCCCC
GCGTCCCCGCGCTCGCCGGCCAGTCAGCTTGCGGGTTCGCTGCCCCGCGAAACCCGAGGTACACAGCCGCGCCTCT
GCTTCCCTGGGCGCGCGCGCCTCCACGCCCTCCTTCTCCCTGGCCCGCGCCTGGCACCAGGGGACCGTTGCCTGA
CGCGAGGCCAGCTCTACTTTTCGCCCCGCGTCTCCTCGCCTGCTGCGCTCTTCCACCAACTCCAACCTCTTCTCCC
TCCAGTCCACTCGCTAGTCCCGACTCCGCCAGCCCTCGGCCCGCTGCGGTAGCGCGCTTCCCGTCCGGTCCCAAA
GGTGGGAACGCGTCCGCCCGCGCCGACCATGGCAGCGTTGCGCTTGGCCGCGCTTCTCTGCACCTGGCAGTGCTC
AGCGCCGCGCTGCTGGCTGCCGAGCTCAAGTCGAAAGTTGCTCGGAAGTGCAGCTCTTACGTGTCCAAAGGCTTC
AACAAGAACGATGCCCCCTCCACGAGATCAACGGTGATCATTGAAGATCTGTCCCAGGGTTCTACCTGCTGCTCT
CAAGAGATGGAGGAGAAGTACAGCTGCAAGTAAAGATGATTCAAAGTGTGGTCAGCGAACAGTGCAATCATTG
CAAGCTGTCTTTGCTTACGTTACAAGAAGTTTGATGAATTCTTCAAAGAACTACTTGAAGTGCAGAGAAATCCCTG
AATGATATGTTGTGAAGACATATGGCCATTTATACATGCAAAATTCTGAGCTATTTAAAGATCTCTTCTAGAGTTG
AAAGTTACTACGTGGTGGGAAATGTGAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCTGGAGCGGATG
TTCGCGCTGGTGAACGCCAGTACCACTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGAGCAGCTGAAG
CCCTTCGGAGATGTCCCTCGCAAATTGAAGTCCAGGTTACTCGTGCTTTTGTAGCAGCCCGTACTTTCTGCTCAAGGC
TTAGCGGTTGCGGGAGATGTCTGAGCAAGGTCTCCGTGGTAAACCCACAGCCAGTGTACCATGCCCTGTTGAAG
ATGATCTACTGCTCCCACTGCCGGGCTCTCGTACTGTGAAGCCATGTTACAATACTGCTCAAACATCATGAGAGGC
TGTTTGGCCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTATAGATGCTATGCTGATGGTGGCAGAGAGGCTA
GAGGTCCTTTCAACATTGAATCGGTATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGGAT
AATAGTGTTCAGGTGTCTCAGAAGGTTTCCAGGGATGTGGACCCCAAGCCCTCCAGCTGGACGAATTTCTCGT
TCCATCTCTGAAAGTGCCTTCAGTGCTCGCTTCCAGCCACATCACCCGAGGAACGCCAACACAGCAGCTGGCACT
AGTTTGGACCGACTGGTTACTGATGTCAAGGAGAACTGAAACAGGCCAAGAAATTTCTGGTCTCCCTTCCGAGCAAC
GTTTGCAACGATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGGAAAGGCAAAAGCAGGTAC
CTGTTTGCACTGACAGGAAATGGATTAGCCAACAGGGCAACAACCCAGAGTCCAGGTTGACACCAGCAACACAGAC
ATACTGATCCTTCGTCAAATCATGGCTCTTCAGTGATGACCAGCAAGATGAAGATGCATACAATGGGAACGACGTG
GACTTCTTTGATATCAGTGATGAAAGTAGTGGAGAAGGAAGTGAAGTGGCTGTGAGTATCAGCAGTGCCCTTCAGAG
TTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGAGAAAGCCGACAGTGCTGGTGCTCCGCTTGGGCA
CAGGCCCTACCTCCTCACTGTCTTCTGCATCTTGTTCCTGGTTATGCAGAGAGAGTGGAGATAAATCTCAAACCTGAG
AAAAAGTGTTCATCAAAAAGTTAAAAGGCACCAAGTTATCACTTTCTACCATCCTAGTGACTTTGCTTTTAAATGAA
TGGACAACAATGTACAGTTTTTACTATGTGGCCACTGGTTTAAAGAGTGTGACTTTGTTTTCTCATTAGTTTTGGG
AGGAAAAGGCACTGTGCATTGAGTTGGTCTGCTCCCCAAACCATGTTAAACGTGGCTAACAGTGTAGGTACAGAA
CTATAGTTAGTTGTGATTTGTGATTTTATCACTCTATTATTTGTTGTATGTTTTTTCTCATTTCGTTTGTGGGTT
TTTTTTTCAACTGTGATCTCGCCTTGTCTTACAAGCAACAGGGTCCCTTCTGGCAGTAACATGTACGTATT
TCTGAAATATTAAATAGCTGTACAGAAGCAGGTTTTATTATCATGTTATCTTATTAAGAAAAAGCCCAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQGST
CCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMFVKTYGH
LYMQNSELFKDLFVELKRYVVGNVNLEEMLNDFWARLLERMFLVNSQYHFTDEYLECVSKYTE
QLKPFQDVPRKLLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHALLKMIYCSHCRGL
VTVKPCYNYCSNIMRGCLANQGDLD FEWNNFIDAMLMVAERLEGPFNIESVMDPIDVKISDAIMN
MQDNSVQVSQKVFQGC GPPKPLPAGRISRSISESAFSARFRPHHPEERPTTAAGTSLDRLVTDVK
EKLKQAKKFWSSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYLFAVTGNGLANQGNNPEVQVDT
KPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGEGSGGCEYQQCPSEFDYNATDHAG
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
 CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAAGCAACT
 TACAGCTGCACCGACAGTTGCGAATGAAAGTTCTAATCTCTTCCCTCCTCCTGTTGCTGCCACTAA
 TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCACAGGGACCGAGGC
 CAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAAAGATTGGTTCTGAG
 AGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAAGAAGCAGTGCCCCCTGTGATCATT
 TCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAAACAAGCATTCCAGA
 GCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTTTGCTCTGCCTTTGTAGGAGCTCTG
 AGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAGACACTC
 TTCTTCTCCACCTCACTCTCCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAAGCA
 TGTTTTTCAAGATCATTTTGTGTTGTTGCTCTCTCTAGTGTCTTCTCTCTCGTCAGTCTTAGCCT
 GTGCCCTCCCCTTACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAACTGTAGCTTCCT
 AGCTAGTGTCAATTAACCTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTT
 AAATGTCAAAAAAAAAAAAAAAAAA

FIGURE 26

MKVLISLLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM
TVSGLPKKQCPCDHFKNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCTCAG
 AGCTGGTCTGCCATGGACATCCTGGTCCCCTCCTGCAGCTGCTGGTGCTGCTTCTTACCCTGCC
 CCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCTGTGCAAAAGCTACTTCCCCTACCTGA
 TGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTCTTCAGCCAG
 ATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGGCTGCGGAACCGGAGC
 CAACTTTTCACTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACCCAAATCCCCACTTTGAGA
 AGTTCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTTGTGGTGGCTCCT
 GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCTGCACTCTGGTGCTGTG
 CTCTGTGCAGAGCCCAAGGAAGGTCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTGC
 TCTTTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGGCCCTTCATGTGGCAGCAAGTTTTTC
 GAGCCCACCTGGAAACACATTGGGGATGGCTGCTGCCTCACCAGAGAGACCTGGAAGGATCTTGA
 GAACGCCCAGTTCTCCGAAATCCAATGGAACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTG
 GGCCCCACATCATGGGAAAGGCTGTCAAACAATCTTTCCCAAGCTCCAAGGCACTCATTGTCTCC
 TTCCCCAGCCTCCAATTAGAACAAGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCTA
GCAGAATGAGAGAAGACATTATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGC
 AATCTCTAACTTCAATCCCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG
 AAACACTAGGACCCTGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTC
 CCAATGTTGTCCCTTTTCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCTCTGAGGCTACAC
 CCATGCGTCTCTAGGAACTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCCTGAC
 CCTCTCTCCCCACTACCACCTTCTTCCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGG
 ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTTGTTCTCAAATATTTTTTAATAAATAGACGAA
 ACCACG

FIGURE 28

MDILVPLLQLLVLLLTPLPLHLMALLGCWQPLCKSYFPYLMVLTTPKSNRKMESKKRELFSSQIKGL
TGASGKVALLELGCGTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDVVVCTLVLCSVQSPRKVLQEVRRVLRPGGVLEFFWEHVAEPYGSWAFMWQQVFEPTW
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPVGPHEMGKAVKQSFSSKALICSFPSL
QLEQATHQPIYLPLRGT

Important features:**Signal peptide:**

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATGTTTGCCTATCCACCTCCCCAAGCCCCTTTACCTATGCTGCTGCTAACGCTGCTGCTGCT
GCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCAGAAAGTCTCTTCTG
CCACTGACGCCCCATCAGGGATTGGGCCTTCTTTCCCCCTTCCTTTCTGTGTCTCCTGCCTCAT
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGTGGGGGATGGCTA
AGAAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGGGCTGCCTTATTTAAA
GTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGCCCTGTTTATTAAGAAATT
GTTCCCTTCCCCTGTGTTCAATGTTTGTAAAGATTGTTCTGTGTAAATATGTCTTTATAATAAAC
AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MLLLTLLLLLLLLLKGSCLWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTCSQAQPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACGTGTGCCAGGCTACCAAGTT
 CCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT
 ATTTGCATCTGTTTTGATAAATGATGTTGACACCTCCACCGAATCTAAGTGGAATCATGTCGG
 GAAGAGATAACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCCATGATGTTTACC
 TTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTTATTGGTTATTTTGGGATTGTTGTT
 TGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACCTCAGCATAGAATTGGACA
 CAGAAAGGGAAAAATATGAAGTGCCTGCTGGGGTTTGCTATCGTATCCACAGGCATCACGGCAGTG
 CTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTTTCCAAATCAC
 AAATAAAGCCATCAGCAGTGCTCCCTTCTGCTGTTCCAGCCACTGTGGACATTTGCCATCCTCA
 TTTTCTTCTGGGTCTCTGGGTGGCTGTGCTGCTGAGCCTGGGAACCTGCAGGAGCTGCCCAGGTT
 ATGGAAGGCGGCCAAGTGAATATAAGCCCTTTTCGGGCATTTCGGTACATGTGGTCGTACCATTT
 AATTGGCCTCATCTGGACTAGTGAATTCATCCTTGGCTGCCAGCAAACTGACTATAGCTGGGGCAG
 TGGTTACTTGTTATTTCAACAGAAGTAAAAATGATCCTCCTGATCATCCCATCCTTTTCGTCTCTC
 TCCATTCTCTTCTTACCATCAAGGAACCGTTGTGAAAGGGTCATTTTTAATCTCTGTGGTGAG
 GATTCCGAGAATCATTGTCATGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCAATGT
 CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTC
 AACCAGAATGCATATACTACAACCTGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGC
 ATTCAAAATCTTGTCCAAGAACTCAAGTCACTTTACATCTATTAAGTCTTTGGAGACTTCATAA
 TTTTCTAGGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTGGAGGACTCATGGCTTTTAACCTAC
 AATCGGGCATTCCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTGCCTACTTAGTAGC
 CCATAGTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCACTTTTCTGTGTTTTGCTGTTGATC
 TGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTTCGTA
 AAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAAGGAATGAGGA
 GGGAACAGAACTCCAGGCCATTGTGAGATAGATACCCATTTAGGTATCTGTACCTGGAAAACATT
 TCCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATTTTTTT
 TTAAAGACCTAATAAACCTATTCTTCCTCAAAA

FIGURE 32

MSGRDTILGLCILALALSLAMMFTFRFITTLVHIFISLVILGLLFVCGVLWWLYDYTNDLSIE
 LDTERENMKCVLGFAIVSTGITAVLLVLI FVLRKRIKLTVELFQITNKAISSAPFLLFQPLWTFA
 ILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQOMTIA
 GAVVTCYFNRSKNDDPPDHPILSSLSILFFYHQGT VVKGSFLISVVRIPRIIVMYMQNALKEQQHG
 ALSRYLFRCCYCCFWCLDKYLLHLNQNAYTTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFGD
 FIIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAHSFLSVFETVLDALFLCFA
 VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33

GTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAGGTTCTTGGACCTCTCCCTGTTTCTTCCTTAGA
 ATAAATTTGTATGGGATTTGTGATGCAGGAAAGCCTAAGGGAAAAAGAATATTCTTGTGTGGT
 GAAAATTTTTTGAAAAAAATTCGCTTCTTCAAACAAGGGTGTCTATCTGATATTTATGAGGAC
 TGTGTCTCTACTATGAAGGCATCTGTTATTGAAATGTTCTTGTGTTTGTGGTGAAGTGGAGTAC
 ATTCAAACAAAGAACGGCAAAGAAGATTAAAAGGCCAAGTTCAGTGTGCCTCAGATCAACTGC
 GATGTCAAAGCCGGAAGATCATCGATCTGAGTTCATTGTGAAATGTCCAGCAGGATGCCAAGA
 CCCCAAATACCATGTTTATGGCACTGACGTGTATGCATCCTACTCCAGTGTGTGTGGCGCTGCCG
 TACACAGTGGTGTGCTTGATAATTCAGGAGGGAAAATACTTGTTCGGAAGGTTGCTGGACAGTCT
 GGTACAAAGGGAGTTATTCCAACGGTGTCCAATCGTTATCCCTACCACGATGGAGAGAATCCTT
 TATCGTCTTAGAAAGTAAACCCAAAAAGGGTGTAACCTACCCATCAGCTCTTACATACTCATCAT
 CGAAAAGTCCAGCTGCCCAAGCAGGTGAGACCACAAAAGCCTATCAGAGGCCACCTATTCCAGGG
 ACAACTGCACAGCCGGTCACTCTGATGCAGCTTCTGGCTGTCACTGTAGCTGTGGCCACCCCCAC
 CACCTTGCCAAGGCCATCCCCCTCTGCTGCTTCTACCACCAGCATCCCCAGACCACAATCAGTGG
 GCCACAGGAGCCAGGAGATGGATCTCTGGTCCACTGCCACCTACACAAGCAGCCAAAACAGGCC
 AGAGCTGATCCAGGTATCCAAAGGCAAGATCCTTCAGGAGCTGCCTTCCAGAAACCTGTTGGAGC
 GGATGTGAGCTGGGACTTGTTCAAAAGAAGAATTGAGCACACAGTCTTTGGAGCCAGTATCCC
 TGGGAGATCCAAACTGCAAAATTGACTTGTGCTTTTTAATTGATGGGAGCACCAGCATTGGCAAA
 CGGCGATTCCGAATCCAGAAGCAGCTCCTGGCTGATGTTGCCAAGCTCTTGACATTGGCCCTGC
 CGGTCCACTGATGGGTGTTGTCCAGTATGGAGACAACCCCTGCTACTCACTTTAACCTCAAGACAC
 ACACGAATTCTCGAGATCTGAAGACAGCCATAGAGAAAATTACTCAGAGAGGAGGACTTTCTAAT
 GTAGGTTCGGGCCATCTCCTTTGTGACCAGAAGTCTTTTCCAAAGCCAATGGAACAGAAAGCGG
 GGCTCCCAATGTGGTGGTGGTGTGATGGTGGATGGCTGGCCACGGACAAAAGTGGAGGAGGCTTCAA
 GACTTGGAGAGAGTCAGGAATCAACATTTTCTTCATACCAATTGAAGGTGCTGCTGAAAATGAG
 AAGCAGTATGTGGTGGAGCCCACTTTGCAAAACAAGGCCGTGTGCAGAACAAACGGCTTCTACTC
 GCTCCACGTGCAGAGCTGGTTTGGCCTCCACAAGACCTGCAGCCTCTGGTGAAGCGGGTCTGCG
 ACACTGACCGCTGGCCTGCAGCAAGACCTGCTTGAACCTGGCTGACATTGGCTTCGTCATCGAC
 GGCTCCAGCAGTGTGGGACGGGCAACTTCCGCACCGTCTCCTCAGTTTGTGACCAACCTCACCAA
 AGAGTTTGAGATTTCCGACACGGACACGCGCATCGGGGCCGTGCAGTACACCTACGAACAGCGGC
 TGGAGTTTGGGTTTCGACAAGTACAGCAGCAAGCCTGACATCCTCAACGCCATCAAGAGGGTGGGC
 TACTGGAGTGGTGGCACCAGCACGGGGGCTGCCATCACTTCGCCCTGGAGCAGCTCTTCAAGAA
 GTCCAAGCCCAACAAGAGGAAGTTAATGATCCTCATCACCGACGGGAGGTCCTACGACGACGTCC
 GGATCCCGAGCCATGGCTGCCCATCTGAAGGGAGTGATCACTATGCGATAGGCGTTGCCCTGGGCT
 GCCCAAGAGGAGCTAGAAGTCATTGCCACTCACCCCGCCAGAGACCACTCCTTCTTGTGGACGA
 GTTTGACAACCTCCATCAGTATGTCCCAGGATCATCCAGAACATTGTACAGAGTTCAACTCAC
 AGCCTCGGAATGAATTTCAGAGCAGGCAGAGCACCAGCAAGTGTGCTTTACTAACTGACGTGTT
 GGACCACCCACCGCTTAATGGGGCACGCACGGTGCATCAAGTCTTGGGCAGGGCATGGAGAAAC
 AAATGTCTTGTATTATTCTTTGCCATCATGCTTTTTTCATATTCCAAAACCTGGAGTTACAAAGA
 TGATCACAAACGTATAGAATGAGCCAAAAGGCTACATCATGTTGAGGGTGTGGAGATTTTACAT
 TTTGACAATTGTTTTCAAATAAATGTTGCGGAATACAGTGCAGCCCTTACGACAGGCTTACGTAG
 AGCTTTTGTGAGATTTTAAAGTTGTTATTCTGATTGAACTCTGTAACCTCAGCAAGTTTCAT
 TTTTGTGATGACAATGTAGGAATTGCTGAATTAATGTTTAGAAGGATGAAAAATAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG
 CQDPKYHVGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYNGVQSLSLPRWR
 ESFIVLESKPKKGVITYPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQLLAVTVAVA
 TPTTLPRPSPSAASTTSIPRPQSVGHRSEQEMDLWSTATYTTSSQNRPRADPGIQRQDPSGAAFQKP
 VGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFRIQKQLLADVAQALDI
 GPAGPLMGVVQYGDNPATHFNLKTHHTNSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
 RSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG
 FYSLHVQSWFGLHKTQLPLVKRVCDTDRILACSKTCLNSADIGFVIDGSSSVGTGNFRTVLQFVTN
 LTKEFEISDTRIGAVQYTYEQRLGFGFDKYSSKPDILNAIKRVGYWSGGTSTGAAINFALEQL
 FKKS KPNKRKLMILITDGRSYDDVRIPAMAAHLKGVITYAIGVAWAAQEELEVIATHPARHDSFF
 VDEFNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
 431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGAAGAAATTGC
 CAAACCATGTCTTTTTTCTGTTTTTCAGAGTAGTTCACAACAGATCTGAGTGTTTTAATTAAGCATGGAAT
 ACAGAAAACAACAAAAAAGCTTAAGCTTTAATTCATCTGGAATCCACAGTTTTCTTAGCTCCCTGGACCC
 GGTGACCTGTTGGCTCTTCCCGCTGGCTGCTCTATCACGTGGTGCTCTCCGACTACTACCCCGAGTGTA
 AAGAACCTTCGGCTCGCGTGCTTCTGAGCTGCTGTGGATGGCCTCGGCTCTCTGGACTGTCTTCCGAGTA
 GGATGTCACTGAGATCCCTCAAATGGAGCCTCCTGCTGCTGCTCACTCCTGAGTTTCTTTGTGATGTGGTAC
 CTCAGCCTTCCCACTACAATGTGATAGAACGCGTGAAGTGGATGTACTTCTATGAGTATGAGCCGATTTA
 CAGACAAGACTTTCACCTTCACACTTCGAGAGCATTCAAAGTCTCTCATCAAAATCCATTTCTGGTCATTCT
 TGGTGACCTCCCACTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTACTTGGGGTAAAAAAGTCT
 TGGTGGGGATATGAGGTTCTTACATTTTCTTATTAGGCCAAGAGGCTGAAAGGAAGACAAAATGTTGGC
 ATTGTCTTAGAGGATGAACACCTTCTTTATGGTGACATAATCCGACAAGATTTTTTAGACACATATAATA
 ACCTGACCTTGAAAACCATATGGCATTGAGGTGGGTAACTGAGTTTGCCCCAATGCCAAGTACGTAATG
 AAGACAGACACTGATGTTTTCACTCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGA
 GAAGTTTTTCAGAGTTATCCTCTAATTGATAATTATCCTATAGAGGATTTTACCAAAAACCCATATTT
 CTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTGGGTTATATAATGTCCAGAGAT
 TTGGTGCCAAGGATCTATGAAATGATGGGTACGTAAAACCCATCAAGTTGAAGATGTTTATGTCCGGAT
 CTGTTTGAATTTATTAAAAGTGAACATTCATATTCAGAAGACACAAATCTTTTCTTCTATATAGAATCC
 ATTTGGATGTCTGCAACTGAGACGTGTGATTGCAGCCCATGGCTTTTCTTCCAAGGAGATCATCACTTTT
 TGGCAGGTCATGCTAAGGAACACCACATGCCATTATTAAGTTCACATTCTACAAAAGCCTAGAAGGACAG
 GATACCTTGTGGAAAGTGTTAAATAAAGTAGGTACTGTGGAAAATTCATGGGGAGGTGAGTGTGCTGGCTT
 AACTGAACTGAACTCATGAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTATTAGTCAGG
 CCCTTCAAAGATGATATGTGGAGGAATTAAATATAAAGGAATTGGAGGTTTTTGCTAAAGAAATTAATAGG
 ACCAAACAATTTGGACATGTCATTCTGTAGACTAGAATTTCTTAAAAGGTGTTACTGAGTTATAAGCTCA
 CTAGGCTGTAAAAACAAAACAAATGTAGAGTTTATTTATTGAACAATGTAGTCACTTGAAGGTTTTGTGTA
 TATCTTATGTGGATTACCAATTTAAAAATATATGTAGTTCTGTGTCAAAAACTTCTTCACTGAAGTTATA
 CTGAACAAAATTTTACCTGTTTTTGGTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTACAGTTATT
 ATTATTTAAATTAATCACTTTGTGTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAGGATAG
 TGAATCATTCCTTACATGCAACATTTCCAGTACTTAACTGATCAGTTTATTATTGATACATCACTCCA
 TTAATGTAAAGTCATAGGTCATTATTGCATATCAGTAATCTCTTGGACTTTGTTAAATATTTTACTGTGGT
 AATATAGAGAAGAATTAAAGCAAGAAATCTGAAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPYRQDFHF
TLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWSWGVEVLTFFLLGQAEKEDKMLA
LSLEDEHLLYGDIIRQDFLDTYNNLTLKTIMAFRWVTEFCPNAKYVMKTDTDVFINTGNLVKYL
NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCSGLGYIMSRDLVPRIYEMMGHV
KPIKFEDVYVGICLNLKVNIHIPEDTNLFFLYRIHLDVCQLRRVIAAHGFSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

CGCTCGGGCACCAGCCGCGGCAAGGATGGAGCTGGGTGCTGGACGCAGTTGGGGCTCACTTTTCTTCAGCTCCTTCTCATC
TCGTCCTTGCCAAGAGAGTACACAGTCATTAATGAAGCCTGCCCTGGAGCAGAGTGGAAATATCATGTGTCGGGAGTGCTGTG
AATATGATCAGATTGAGTGCGTCTGCCCCGGAAGAGGGAAGTCGTGGGTATACCATCCCTTGCTGCAGGAATGAGGAGAA
TGAGTGTGACTCCTGCTGATCCACCCAGGTTGTACCATCTTTGAAAAGTGCAGAGCTGCCGAAATGGCTCATGGGGGGT
ACCTTGGATGACTTCTATGTGAAGGGGTCTACTGTGCAGAGTGCCGAGCAGGCTGGTACGGAGGAGACTGCATGCGATGTG
GCCAGGTTCTGCGAGCCCCAAGGGTCAGATTTTGTGGAAAGCTATCCCTAAATGCTCACTGTGAATGGACCATTTCATGC
TAAACCTGGGTTTGTCTATCCAACATAAGATTTGTCTGTGAGTCTGGAGTTTACTACATGTGCCAGTATGACTATGTTGAG
GTTCTGTATGGAGACAACCCGATGGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGGCCAGCTCCTATCCAGAGCATAG
GATCCTCACTCCACGTCCTCTTCCACTCCGATGGCTCCAAGAATTTTACGGTTTCCATGCCATTTATGAGGAGATCACAGC
ATGTCCTCTATCCCTTGTTCATGACGGCACGTGCGTCCTTGACAAGGCTGGATCTTACAAGTGTGCTGCTTGGCAGGC
TATACTGGGCAGCGCTGTGAAAATCTCCTTGAAGAAAGAACTGCTCAGACCCTGGGGGGCCAGTCAATGGGTACCAGAAAA
TAACAGGGGGGGCTGGGCTTATCAACGGACGCCATGCTAAATTTGGCACCCTGGTGTCTTTCTTTTGTAACTACCTATGT
TCTTAGTGGCAATGAGAAAAGAACTTGCCAGCAGAAATGGAGAGTGGTCAGGGAACAGCCCATCTGCATAAAAGCTGCCGA
GAACCAAGATTTTCAAGCCTGGTGAGAAAGGAGATTCTTCCGATGCAGGTTTCAAGGAGACACCATTAACACAGCTAT
ACTCAGCGGCTTTCAGCAAGCAGAACTGCAGAGTGGCCCTACCAAGAAGCCAGCCCTTCCCTTTGGAGATCTGCCATGGG
ATACCAACATCTGCATACCCAGCTCCAGTATGAGTGCATCTACCCCTTCTACCGCCGCTGGGCAGCAGCAGGAGGACATGT
CTGAGGACTGGGAAGTGGAGTGGGCGGGCACCATCTGCATCCCTATCTGCGGGAAAATTGAGAACATCACTGCTCCAAAGA
CCCCAAGGGTTGCGCTGGCCGTGGCAGGCAGCCATCTACAGGAGGACCAGCGGGGTGCATGACGGCAGCCTACACAAGGGAGC
GTGGTTCTTAGTCTGCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGTGGCTGCCCACTGTGTACTGACCTGGGGAAG
GTCACCATGATCAAGACAGCAGACTGAAAAGTTGTTTGGGGAAATTTACCGGGATGATGACCGGGATGAGAAGACCATCC
AGAGCCTACAGATTTCTGCTATCATTTCTGCATCCCAACTATGACCCCATCTGCTTGATGCTGACATCGCCATCTGAAGCT
CCTAGACAAGGCCGATATCAGCACCAGTCCAGCCATCTGCCCTCGCTGCCAGTCGGGATCTCAGCACTTCTTCCAGGAG
TCCCACATCACTGTGGCTGGCTGGAATGTCTGGCAGACGTGAGGAGCCCTGGCTTCAAGAACGACACACTGCGCTCTGGGG
TGGTCAGTGTGGTGAGCTCGCTGCTGTGTGAGGAGCAGCATGAGGACCATGGCATCCCACTGAGTGTCACTGATAACATGTT
CTGTGCCAGCTGGGAACCACTGCCCTTCTGATATCTGCACTGCAGAGACAGGAGGCATCGCGGCTGTGTCTTCCCGGGA
CGAGCATCTCCTGAGCCACGCTGGCATCTGATGGGACTGGTCAGCTGGAGCTATGATAAAACATGCAGCCACAGGCTCTCCA
CTGCCCTTACCAAGGTGCTGCCCTTTAAAGACTGGATTGAAAGAAATATGAAATGAACCATGCTCATGCATCCTTGAGAAG
TGTTTCTGTATATCCGCTGTGTACGTGTGTCATTGCGTGAAGCAGTGTGGGCTGAAGTGTGATTTGGCCTGTGAATTTGGCT
GTGCCAGGGCTTCTGACTTCAGGGACAAAACCTCAGTGAAGGTTGAGTAGACCTCCATTGCTGGTAGGCTGATGCCGCTCCA
CTACTAGGACAGCCAATTGGAAGATGCCAGGGCTTGCAAGAAGTAAGTTTCTTCAAGAAGACCATATACAAAACCTCTCCA
CTCCACTGACCTGGTGGTCTTCCCCAACTTTCACTTATACGAATGCCATCAGCTTGACCAGGGAAGATCTGGGCTTCATGAG
GCCCTTTTGGGCTCTCAAGTTCTAGAGAGCTGCCCTGTGGGACAGCCAGGGCAGCAGAGCTGGGATGTGGTGCATGCCCTT
TGTGTACATGGCCACAGTACAGTCTGGTCCCTTTTCTTCCCCATCTCTGTACACATTTAATAAAATAAGGGTTGGCTCT
GAAC TACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 38

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVVGYT
 IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGLDDFYVKGFYCAECRAGWYGGDCMRCGQ
 VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDQII
 KRVCGNERPAPIQSIGSSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC
 ACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNSYVLSGNE
 KRTCQQNGEWSGKQPICIKACREPKISDLVRRRVLPMPVQSRETPHQLYSAAFSKQKLQSAPTK
 KPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRRTCLRTGKWSGRAPSCIPICGKIENITAP
 KTQGLRWPWQAAYRRTSGVHDGSLHKGAWFLVCSGALVNERTVVVAHCVTDLGKVTMIKTADL
 KVVLGKFYRDDDRDEKTIQSLQISAILHPNYDPILLDADIAILKLLDKARISTRVQPICLAASR
 DLSTSFQESHITVAGWNVLADVRS PGFKNDTLRSGVSVVDSLLCEEQHEDHGIPVSVTDNMFCA
 SWEPTAPSDICTAETGGIAAVSFPGRASPEPRWHLMGLVSWSYDKTCSHRLSTAFTKVLFPKDWI
 ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
 474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTTCCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCGTGATTTATTAACTGGCTTAATC
TGAAGGTTCTCAGTCAAATCTTTGTGATCTACTGATTGTGGGGGCATGGCAAGTTTGCTTAAAGGAGCTTGGCTGG
TTTGGGCCCTTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGAAATGAAGGCGCTTCTGTTGC
TGGTCTTGCCCTGGCTCAGTCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCTGTATTGAGAACTCTGTA
AAGGTGCTCCCACTACGGCCTGACCAAAGATAGGAAGAGGCGCTCACAAGATGGCTGTCCAGACGGCTGTGCGAGCC
TCACAGCCACGGCTCCCTCCCAAGAGGTTTCTGCAGCTGCCACCATCTCTTAATGACAGACGAGCTTGGCTTAGACA
ACCTTGCTACGTGTCTCGGCAGAGGACGGGCGAGCCAGCAATCAGCCAGTGGACTCTGGCCGGAGCAACCGAACTA
GGGCAGGCGCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAAAAAATAAATCGAGCTTTGAGTGTTCTTCGAA
GGACAAAGAGCGGGAGTGCACTTGCCAACCATGCCGACCGGAGGAGGAAAATTCTGAAACACCACTGCCCTGAAG
TCTTTCCAAGGTTGTACCACTGATTCCAGATGGTGAAATACCAGCATCAAGATCAATCGAGTAGATCCAGTGAA
GCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAACCCCACTGGTCCATATCATTATCCAACACATTTATCGTGATGGG
TGATCGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTC
ACAACTACGCTGTGCTCTCTGCGGCAGCCCTGCCAGGTGCTGTGGCTGACTGTGATGCGTGAACAGAAGTCCGCA
GCAGGAACAATGGACAGGCCCCGGATGCTACAGACCCCGAGATGACAGCTTTCATGTGATTCTCAACAAAAGTAGCC
CCGAGGAGCAGCTTGAATAAACTGGTGGCAAGGTGGATGAGCCTGGGGTTTTTCATCTTCAATGTGCTGGATGGCG
GTGTGGCATATCGACATGGTCAGCTTGAGGAGAATGACCGTGTGTAGCCATCAATGGACATGATCTTCGATATGGCA
GCCCAGAAAGTGCGGCTCATCTGATTCAAGCCAGTGAAAGACGTGTTACCTCGTCGTGTCCCGCCAGGTTCCGCGAGC
GGAGCCCTGACATCTTTCAGGAAGCGCGCTGGAACAGCAATGGCAGCTGGTCCCGAGGCGAGGGGAGAGGAGCAACA
CTCCCAAGCCCCCTCCATCTTACAATTACTTGTCTGATGAGAAGGTGGTAAATATCCAAAAAGACCCCGTGAATCTCTCG
GCATGACCGTCGAGGGGGAGCATCACATAGAGAATGGGATTTGCCTATCTATGTCTCATCAGTGTTGAGCCCGGAGGAG
TCATAAGCAGAGATGGAAGAATAAAACAGGTGACATTTTGTGAATGTGGATGGGGTCGAAGTACAGAGGTCAGCC
GGAGTGAGGCACTGGCATTATTGAAAAGAACATCATCCTCGATAGTACTCAAAGCTTTGGAAGTCAAAGAGTATGAGC
CCCAGGAAGACTGCAGCAGCCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCCAGTGACTGGTCCCCATCCT
GGGTGATGTGGCTGGAATTACCAAGGTGCTTGTATAACTGTAAAGATATTGTATTACGAAGAAACACAGCTGGAAGTC
CACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATCTTCTTGCTGTCAATGGTAGAAGTACATCAGGAATGA
TACATGCTTGTGCTGGCAAGACTGCTGAAAGAACTTAAAGGAAGAATTACTCTAACTATTGTTTCTTGGCCTGGCACTT
TTTTATAGAATCAATGATGGGTGAGAGGAAAACAGAAAAATCACAAATAGGCTAAGAAGTTGAAACACTATATTTATC
TTGTGAGTTTTTATATTTAAAGAAAGAAATACATTGTAAAAATGTCAGGAAAAGTATGATCATCTAATGAAAGCCAGTT
ACACCTCAGAAAAATATGATTCAAAAAAATTAACCTACTAGTTTTTTTTCAGTGTGGAGGATTTCTCATTAATCTAC
AACATGTTTATATTTTCTATTCAATAAAAGCCCTAAACAACTAAATGATTGATTGTATACCCCACTGAATT
CAAGCTGATTAAATTTAAATTTGGTATATGCTGAAGTCTGCCAAGGGTACATTATGGCCATTTTAAATTTACAGCT
AAAAATTTTTTAAATGCATTGCTGAGAAACGTGCTTTCATCAACAAGAATAAATATTTTTTCAGAAGTTAA

FIGURE 40

MKALLLLVLPWLS PANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPCDGCASLTATAPS
 PEVSAAATISLMTDEPLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRSFKKINR
 ALSVLRRTKSGSAVANHADQGRESENTTAPEVFPRLYHLIPDGEITSIKINRVPSESLSIRLV
 GGSETPLVHIIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLLRQPCQVLWLTVM
 REQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVKRVDEPGVFI FNVLDGGVAYRHG
 QLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRQRSPDIFQEAGWNSNGSWSPG
 PGERSTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHREWDLPIYVISVEPGGVISRDR
 IKTGDIILLNVGDVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAP
 SDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFIFKSIVEGTPAYNDG
 RIRCGDILLAVNGRSTSGMIHACLARLLKELKGRITLTIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
 467-473, 603-609

FIGURE 41

ACCAGGCATTGTATCTTCAGTTGTTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGAAGCTTT
 CTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTCACGTAATAAAAAACATGGGCTTCAACCTGACT
 TTCCACCTTTTCCTACAAATCCGATTACTGTTGCTGTTGACTTTGTGCCTGACAGTGGTTGGGTGGGC
 CACCAGTAACTACTTCGTGGGTGCCATTCAAGAGATTCCCTAAAGCAAAGGAGTTCATGGCTAATTTCC
 ATAAGACCCCTCATTTTGGGAAGGGAAAACTCTGACTAATGAAGCATCCACGAAGAAGGTAGAACTT
 GACAACTGTCTTCTGTGTCTCCTTACCTCAGAGGCCAGAGCAAGCTCATTTTCAAACCAGATCTCAC
 TTTGGAAGAGGTACAGGCAGAAAATCCCAAAGTGTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAAG
 CTTTACAGAGGGTCGCCATCCTCGTTCCCCACCGAACAGAGAGAAACACCTGATGTACCTGCTGGAA
 CATCTGCATCCCTTCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGG
 TAAAAAGTTTAATCGAGCCAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGG
 ACTGCTTTATATCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGAGGAG
 CATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTGGATATTTTGG
 GGGTGTTACTGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCTAACAACCTACTGGGGAT
 GGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTTCCCGGCCCTG
 CCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACG
 GATGAAGCTCTTACACCAAGTGTACGAGTCTGGAGAACAGATGGGTGAGTAGTTGTTCTTATAAAT
 TAGTATCTGTGGAACACAATCCTTTATATATCAACATCACAGTGGATTCTGGTTTGGTGCATGACCC
 TGGATCTTTTGGTGATGTTTGAAGAAGTGAATCTTTGTTTGAATAATTTTGGCCTAGAGACTTCAA
 ATAGTAGCACACATTAAGAACCTGTTACAGCTCATTTGTTGAGCTGAATTTTCTTTTGTATTTTCT
 TAGCAGAGCTCCTGGTGATGTAGAGTATAAACAGTTGTAACAAGACAGCTTTCTTAGTCATTTTGT
 CATGAGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAGGATGACTCAAAGGATAA
 AATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATATTATGGGAT
 AAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCTCGTCCAAGGTAGAA
 AGGTACGAAGATACAATACTGTTATTCAATTTATCCTGTACAATCATCTGTGAAGTGGTGGTGCAGGT
 GAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCAGGACACAGTGAACCTGGGAATGAAGA
 GGTAGCAGGAGGGTGGAGTGTGGCTGCAAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGATAGC
 CTTCAAGGGAGGACCTGCCAGGTATGCCCTTCCAGTGATGCCACCAGAGAATACATTCTCTATTAGT
 TTTTAAAGAGTTTTTGTAAATGATTTTGTACAAGTAGGATATGAATTAGCAGTTTACAAGTTTACAT
 ATTAATAATAAATATGTCTATCAAATACCTCTGTAGTAAAATGTGAAAAAGCAAAA

FIGURE 42

MGFNLTFFHLSYKFRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTILGKGKTLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKKFNRAKLLNVGYLEALKEENWDCFI FHDV
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGVTALSREQFFKVNGFSNNYWGWWGED
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 44

MALSSQIWAACLLLLLLSLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTHFPI
CIFCCGCCHRSKCGMCCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCCAACATGCCTCACCC
TCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT
TCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTCTATTGTCTG
GACCTTCAACACAACCCCTCTTGTCAACATACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAACTGAAG
AAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCA
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATT
TATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTC
CTGGAGATGGGGAGAAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCTGTCAGCAGAACT
TCTCAAGCCCCATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCCCTCCATG
GTCCTCCTGTGTCTCCTGTTGGTGCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTG
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGACATTTGTCTGGG
AAACTCCTAACATATGCCCCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
AGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGAT
GGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTA
TCTTAGACAGCAGTGCCTCCCTAAGTCTCTGCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVTIQP
 EGGTIIIVTQNRNRERVDFFDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
 PKVTMGLQSNKNGTCVTNLTCMEHGEEVDIYTWKALGQAANESHNGSILPISWRWGESDMTFIC
 VARNPVSRNFSSPILARKLCEGAADDPSSMVLCLLLVPLLLSLFVLGLFLWFLKRERQEEYIE
 EKKRVDICRETPNICPHSGENTYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLTMPDT
 PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
 291-295

FIGURE 47

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGG
ATTGAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTCAGCT
TAGTTGAGGAAGACCAATTTTCTCAAAACCCCATCTCTTGCTTTGAGTGGTGGTCCCAGGAATT
ATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAAAGAGCGTG
CTGCAACAACAGAACTGGAATGTTTCTTTCATCATTTTTCAGTGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCCTCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGACATTCATCCAGAATCCTTCAA
CTTGCAAGTGGTTTTTCAATGACTCTTGTGCACCTCCTACTGGTTTCAATAAACCCACCAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCACTTCGATTCTGAAGAAAACAAACATAGG
CTTATCCACTTCTCAGTATTTTTAGGTCTATTGCTTGTGGAATTCTGGAGGTCCTGTTTGGGCT
CAGTCAGATAGTCATCGGTTTCTTGGCTGTCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG
TGTAGTTTAAATGGGAATAAAATGTAAGTATCAGTAGTTTGAAAAAAAAA

FIGURE 48

MTCCEGWTSCNGFSLVLLVLLLVGVVLAIPVIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
 TTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCLISIQALLKGPLMCNSPSNSNANCEFSL
 KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSSEENKHRLIHFSVFLGL
 LLVGILEVLFLGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGGTGA
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
ATCCGTGGGCTGCAGACCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCCCTCGAACTGTGA
CATGGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGGAAGCCAATGACC
CATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
TGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGGCAAATGCAAATACAAGAG
CAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCCACTCATCACTCCAGGCTCTGCCA
CTACTTGCTGAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCAGCACC
TCCTCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCCCT
TTCTGATCAGGAGGCTTCTTTATGAATTAACTCGCCCCACCACCCCTCA

FIGURE 50

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAAGIAAVLSGKCKYKS
SQKQHSPVPEKAIPILITPGSATTC

Important features:**Signal peptide:**

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG
AGGACAGGGAGTCGGAAGGAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGCAAGG
AGGAGACCTTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGGCTGGGCAGAGATGAAGTTCCAG
GGGCCCCCTGGCCTGCCTCCTGCTGGCCCTCTGCCTGGGCAGTGGGGAGGCTGGCCCCCTGCAGAG
CGGAGAGGAAAGCACTGGGACAAATATTGGGGAGGCCCTTGGACATGGCCTGGGAGACGCCCTGA
GCGAAGGGGTGGGAAAGGCCATTGGCAAAGAGGCCGGAGGGGGCAGCTGGCTCTAAAGTCAGTGAG
GCCCTTGGCCAAGGGACCAGAGAAGCAGTTGGCACTGGAGTCAGGCAGGTTCCAGGCTTTGGCGC
AGCAGATGCTTTGGGCAACAGGGTCGGGGAAGCAGCCCATGCTCTGGGAAACACTGGGCACGAGA
TTGGCAGACAGGCAGAAGATGTCAATTCGACACGGAGCAGATGCTGTCCGCGGCTCCTGGCAGGGG
GTGCCTGGCCACAGTGGTGCTTGGGAACTTCTGGAGGCCATGGCATCTTTGGCTCTCAAGGTGG
CCTTGGAGGCCAGGGCCAGGGCAATCCTGGAGGTCTGGGGACTCCGTGGGTCCACGGATACCCCG
GAAACTCAGCAGGCAGCTTTGGAATGAATCCTCAGGGAGCTCCCTGGGGTCAAGGAGGCAATGGA
GGGCCACCAAACCTTTGGGACCAACACTCAGGGAGCTGTGGCCAGCCTGGCTATGGTTCAGTGAG
AGCCAGCAACCAGAATGAAGGGTGACGAATCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA
ACTCTGGGGGAGGCAGCGGCTCACAGTCGGGCAGCAGTGGCAGTGGCAGCAATGGTGACAACAAC
AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCGGCAGCAG
TGGCGGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGGCAGTGAGT
CCTCCTGGGGATCCAGCACCGGCTCCTCCTCCGGCAACCACGGTGGGAGCGGCGGAGGAAATGGA
CATAAACCCGGGTGTGAAAAGCCAGGGAATGAAGCCCGCGGAGCGGGGAATCTGGGATTCAGGG
CTTCAGAGGACAGGGAGTTTCCAGCAACATGAGGGAATAAGCAAAGAGGGCAATCGCCTCCTTG
GAGGCTCTGGAGACAATTATCGGGGGCAAGGGTCGAGCTGGGGCAGTGGAGGAGGTGACGCTGTT
GGTGGAGTCAATACTGTGAACCTCTGAGACGTCTCCTGGGATGTTTAACTTTGACACTTTCTGGAA
GAATTTTAAATCCAAGCTGGGTTTCATCAACTGGGATGCCATAAACAAGGACCAGAGAAGCTCTC
GCATCCCGTGAACCTCCAGACAAGGAGCCACCAGATTGGATGGGAGCCCCCACACTCCCTCCTTAA
AACACCACCTCTCATCTAATCTCAGCCCTTGCCCTTGAAATAAACCTTAGCTGCCCCACAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

1. NAME _____
 2. DATE _____
 3. TIME _____
 4. LOCATION _____
 5. REASON _____
 6. WITNESSES _____
 7. SIGNATURE _____
 8. INITIALS _____
 9. REMARKS _____
 10. DATE _____
 11. TIME _____
 12. LOCATION _____
 13. REASON _____
 14. WITNESSES _____
 15. SIGNATURE _____
 16. INITIALS _____
 17. REMARKS _____
 18. DATE _____
 19. TIME _____
 20. LOCATION _____
 21. REASON _____
 22. WITNESSES _____
 23. SIGNATURE _____
 24. INITIALS _____
 25. REMARKS _____
 26. DATE _____
 27. TIME _____
 28. LOCATION _____
 29. REASON _____
 30. WITNESSES _____
 31. SIGNATURE _____
 32. INITIALS _____
 33. REMARKS _____
 34. DATE _____
 35. TIME _____
 36. LOCATION _____
 37. REASON _____
 38. WITNESSES _____
 39. SIGNATURE _____
 40. INITIALS _____
 41. REMARKS _____
 42. DATE _____
 43. TIME _____
 44. LOCATION _____
 45. REASON _____
 46. WITNESSES _____
 47. SIGNATURE _____
 48. INITIALS _____
 49. REMARKS _____
 50. DATE _____
 51. TIME _____
 52. LOCATION _____
 53. REASON _____
 54. WITNESSES _____
 55. SIGNATURE _____
 56. INITIALS _____
 57. REMARKS _____
 58. DATE _____
 59. TIME _____
 60. LOCATION _____
 61. REASON _____
 62. WITNESSES _____
 63. SIGNATURE _____
 64. INITIALS _____
 65. REMARKS _____
 66. DATE _____
 67. TIME _____
 68. LOCATION _____
 69. REASON _____
 70. WITNESSES _____
 71. SIGNATURE _____
 72. INITIALS _____
 73. REMARKS _____
 74. DATE _____
 75. TIME _____
 76. LOCATION _____
 77. REASON _____
 78. WITNESSES _____
 79. SIGNATURE _____
 80. INITIALS _____
 81. REMARKS _____
 82. DATE _____
 83. TIME _____
 84. LOCATION _____
 85. REASON _____
 86. WITNESSES _____
 87. SIGNATURE _____
 88. INITIALS _____
 89. REMARKS _____
 90. DATE _____
 91. TIME _____
 92. LOCATION _____
 93. REASON _____
 94. WITNESSES _____
 95. SIGNATURE _____
 96. INITIALS _____
 97. REMARKS _____
 98. DATE _____
 99. TIME _____
 100. LOCATION _____
 101. REASON _____
 102. WITNESSES _____
 103. SIGNATURE _____
 104. INITIALS _____
 105. REMARKS _____
 106. DATE _____
 107. TIME _____
 108. LOCATION _____
 109. REASON _____
 110. WITNESSES _____
 111. SIGNATURE _____
 112. INITIALS _____
 113. REMARKS _____
 114. DATE _____
 115. TIME _____
 116. LOCATION _____
 117. REASON _____
 118. WITNESSES _____
 119. SIGNATURE _____
 120. INITIALS _____
 121. REMARKS _____
 122. DATE _____
 123. TIME _____
 124. LOCATION _____
 125. REASON _____
 126. WITNESSES _____
 127. SIGNATURE _____
 128. INITIALS _____
 129. REMARKS _____
 130. DATE _____
 131. TIME _____
 132. LOCATION _____
 133. REASON _____
 134. WITNESSES _____
 135. SIGNATURE _____
 136. INITIALS _____
 137. REMARKS _____
 138. DATE _____
 139. TIME _____
 140. LOCATION _____
 141. REASON _____
 142. WITNESSES _____
 143. SIGNATURE _____
 144. INITIALS _____
 145. REMARKS _____
 146. DATE _____
 147. TIME _____
 148. LOCATION _____
 149. REASON _____
 150. WITNESSES _____
 151. SIGNATURE _____
 152. INITIALS _____
 153. REMARKS _____
 154. DATE _____
 155. TIME _____
 156. LOCATION _____
 157. REASON _____
 158. WITNESSES _____
 159. SIGNATURE _____
 160. INITIALS _____
 161. REMARKS _____
 162. DATE _____
 163. TIME _____
 164. LOCATION _____
 165. REASON _____
 166. WITNESSES _____
 167. SIGNATURE _____
 168. INITIALS _____
 169. REMARKS _____
 170. DATE _____
 171. TIME _____
 172. LOCATION _____
 173. REASON _____
 174. WITNESSES _____
 175. SIGNATURE _____
 176. INITIALS _____
 177. REMARKS _____
 178. DATE _____
 179. TIME _____
 180. LOCATION _____
 181. REASON _____
 182. WITNESSES _____
 183. SIGNATURE _____
 184. INITIALS _____
 185. REMARKS _____
 186. DATE _____
 187. TIME _____
 188. LOCATION _____
 189. REASON _____
 190. WITNESSES _____
 191. SIGNATURE _____
 192. INITIALS _____
 193. REMARKS _____
 194. DATE _____
 195. TIME _____
 196. LOCATION _____
 197. REASON _____
 198. WITNESSES _____
 199. SIGNATURE _____
 200. INITIALS _____
 201. REMARKS _____
 202. DATE _____
 203. TIME _____
 204. LOCATION _____
 205. REASON _____
 206. WITNESSES _____
 207. SIGNATURE _____
 208. INITIALS _____
 209. REMARKS _____
 210. DATE _____
 211. TIME _____
 212. LOCATION _____
 213. REASON _____
 214. WITNESSES _____
 215. SIGNATURE _____
 216. INITIALS _____
 217. REMARKS _____
 218. DATE _____
 219. TIME _____
 220. LOCATION _____

MKFQGGLACLLALCLSGSEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGAAGSKVS
 EALGQGTREAVGTGVRQVPFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGVF
 GHSGAWETSGGHGIFGSSQGGGLGGQGGNPGGLGTPFWVHGYPGNSAGSFGMNPQGAPWQGGNGGPPNF
 GTNTQGAVAPQGYGSVRASNQNEGCTNPPPSGSGGGSSNSGGSGSGSQSGSSGSGSNGDNNNGSSSSGGS
 SSGSSSSGSSSSGGSSGGSSSGNSGGSRGDSGSESSWGSSTGSSSGNHGGSGGGNGHKGPGCEKPGNE
 ARGSGESGIQGFRRQGVSSNMREISKEGNRLGGSGDNRYRGQSSWGSGGGDVAGGVNTVNSETSPGM
 FNFDTFWKNFKSKLGFINWDAINKDQRSSRI P

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,
90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,
159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,
236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,
253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,
283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,
298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,
325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGGATGTCGCTGCTGAGCCTGCCCTGG
 CTGGGCTCAGACCGGTGGCAATGTCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCTGGCT
 ACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAACCTGCCGCCGGCTCCAGTGTTTCC
 CACAGCCCCAAAACGGAACCTGGTTTTGGGGTCACCTGGGCCTGATCACTCTACAGAGGAGGGC
 TTGAAGGACTCGACCCAGATGTCGGCCACCTATTCAGGGCTTTACGGTATGGCTGGGTCCCAT
 CATCCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCACCATGCCTCAGCTGCCA
 TTGCACCCAAGGATAATCTCTTCATCAGGTTCTGAAGCCCTGGCTGGGAGAAGGGATACTGCTG
 AGTGGCGGTGACAAGTGGAGCGCCACCGTCGGATGCTGACGCCCGCCTTCCATTTCAACATCCT
 GAAGTCTATATAACGATCTTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCAGCACCTGG
 CCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAGCACATCAGCCTCATGACCTTGGACAGTCTA
 CAGAAATGCATCTTCAGCTTTGACAGCCATTGTGAGGAGAGGCCAGTGAATATATTGCCACCAT
 CTTGGAGCTCAGTGCCCTTGAGAGAAAAGAAGCCAGCATATCCTCCAGCACATGGACTTTCTGT
 ATTACCTCTCCCATGACGGGCGGCGCTTCCACAGGGCCTGCCGCTGGTGCATGACTTCACAGAC
 GCTGTATCCGGGAGCGGCGTCGCACCCTCCCCACTCAGGGTATTGATGATTTTTTCAAAGACAA
 AGCCAAGTCCAAGACTTTGGATTTTCATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGG
 CATTGTGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCAG
 GCCAGTGGCCTCTCCTGGGTCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCG
 ACAGGAGGTGCAAGAGCTTCTGAAGGACCGCATCTAAAGAGATTGAATGGGACGACCTGGCCC
 AGCTGCCCTTCCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCAGCTCCCTTCATC
 TCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTG
 CCTCATCGATATTATAGGGGTCCATCACAACCCAACCTGTGTGGCCGGATCCTGAGGTCTACGACC
 CCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTACCTCTGGCTTTTATTCCTTTCTCCGCA
 GGGCCCAGGAACCTGCATCGGGCAGGCGTTCGCCATGGCGGAGATGAAAGTGGTCTGGCGTTGAT
 GCTGCTGCACTTCCGGTTCCTGCCAGACCACACTGAGCCCCGAGGAAGCTGGAATTGATCATGC
 GCGCCGAGGGCGGGCTTTGGCTGCGGGTGGAGCCCCCTGAATGTAGGCTTGAGTGA¹CTTTCTGAC
 CCATCCACCTGTTTTTTTGCAGATTGTTCATGAATAAAACGGTGCTGTCAA

FIGURE 54

MSLLSLPWLGLRPVAMSPWLLLLLVGSWLLARILAWTYAFYNNCRRRLQCFPQPKNWFWGHLG
LITPTEEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP
WLGEIGILLSGGDKWSRHRRLTPAFHFNIKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI
SLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRRFHRAC
RLVHDFTDVIRERRRTLPTQGIDDFKDKAKSKTLDLFDVLLLSKDEDGKALSDEDIRAEADTF
MFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPEVYDPFRFDPENSKGRSP
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN
VGLQ

Important features:**Transmembrane domains:**

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTTTGA
 GCCTACTCGTTTGATTGCAACTATCATGGTGCTGTTGTGTTTTGCACTTACCCTGTGTTCTGCCT
 TTTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTTGCAGTCTTTGGCATTGACGTGG
 TACAGCCTTTCCTTCATACCATTGCAAGGGATGCTGTGAAGAAGTGTTTTGCCGTGTGTCTTGC
ATAATTCATGGCCAGTTTATGAAGCTTTGGAAGGCACTATGGACAGAAGCTGGTGACAGTTT
 GTAACCTATCTTCGAAACCTCTGTCTTACAGACATGTGCCTTTTATCTTGACAGCAATGTGTTGCTT
 GTGATTCGAACATTTGAGGGTTACTTTTGAAGCAACAATACATTCTCGAACCTGAATGTCAGTA
 GCACAGGATGAGAAGTGGGTTCTGTATCTTGTGGAGTGGAATCTTCCTCATGTACCTGTTTCCTC
 TCTGGATGTTGTCCCACTGAATTCCCATGAATACAAACCTATTCAGCAACAGCAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYSLSFIPFAR
DAVKKCFAVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

[illegible]

FIGURE 58

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSI FKL SVFIP SQEFSTYRQWKQKIVQAGDKDL DG
 QLD FEEFVHYLQDHEKKLRLVFKI LD KNDGRIDAQEIMQSLRDLGVKISEQQA EKILKSMDKNG
 TMTIDWNEW RDYHLLHPVENIPEI ILYWKHSTIFDVGENLTVPDEFTVEERQTGMWWRHLVAGGG
 AGAVSRTCTAPLDR LKVL MQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLK IAPESA I K
 FMAYEQIKRLVGS DQETLRIHERLVAGSLAGAI AQSSIYPMEVLKTRMALRKTGQYSGMLDCARR
 ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKN AWLQHYAVNSADPGVFVLLACGTMSSTC
 GQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGA FGLYRGLAPNFMKVIPAVSISYV
 VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 60

MASLGQILFWSIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASRLKNVQLTDAGTYKC
YIITSKGKGNANLEYKTGAFSMPEVNVVDYNASSETLRCEAPRWFPQPTVVWASQVDQGANFSEVS
NTSFELNSENVTMKVSVLYNVTINNTYSCMIENDIAKATGDIKVTSEIKRRSHLQLLNSKASL
CVSSFFAISWALLPLSPYMLK

Important features:**Signal peptide:**

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTCAGAATCACCATGGCCAGCTATCCTTACCGGCAGGGCTGCCCAGGAGCTGCAGGACAAG
 CACCAGGAGCCCCCTCCGGGTAGCTACTACCCTGGACCCCCCAATAGTGGAGGGCAGTATGGTAGT
 GGGCTACCCCCTGGTGGTGGTTATGGGGGTCTGCCCCCTGGAGGGCCTTATGGACCACCAGCTGG
 TGGAGGGCCCTATGGACACCCAATCCTGGGATGTTCCCCTCTGGAACCTCAGGAGGACCATATG
 GCGGTGCAGCTCCCGGGGGCCCCCTATGGTCAGCCACCTCCAAGTTCCTACGGTGGCCAGCAGCCT
 GGGCTTTATGGACAGGGTGGCGCCCCCTCCAATGTGGATCCTGAGGCCTACTCCTGGTTCCAGTC
 GGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
 ATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTTGACAAGACCAAGTCA
 GGCCGCATCGATGTCTACGGCTTCTCAGCCCTGTGGAAATTCATCCAGCAGTGGAAAGAACCTCTT
 CCAGCAGTATGACCGGGACCGCTCGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCC
 AAATGGGCTACAACCTGAGCCCCAGTTCACCCAGCTTCTGGTCTCCCGCTACTGCCACGCTCT
 GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA
 GGCCTTCGGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTTCGAGGACTTCGTCA
 CCATGACAGCTTCTCGGATGCTATGACCCAACCATCTGTGGAGAGTGGAGTGCACCAGGGACCTT
 TCCTGGCTTCTTAGAGTGAGAGAAGTATGTGGACATCTCTTCTTTTCTGTCCCTCTAGAAGAAC
 ATTCTCCCTTGCTTGATGCAACACTGTTCCAAAAGAGGGTGGAGAGTCTTGCATCATAGCCACCA
 AATAGTGAGGACCGGGGCTGAGGCCACACAGATAGGGGCTGATGGAGGAGAGGATAGAAGTTGA
 ATGTCCTGATGGCCATGAGCAGTTGAGTGGCACAGCCTGGCACAGGAGCAGGTCTTGTAATGG
 AGTTAGTGTCAGTCAGCTGAGCTCCACCCTGATGCCAGTGGTGAGTGTTCATCGGCCTGTTACC
 GTTAGTACCTGTGTTCCCTCACCAGGCCATCCTGTCAAACGAGCCCATTTTCTCCAAAGTGGAAT
 CTGACCAAGCATGAGAGAGATCTGTCTATGGGACCAGTGGCTTGGATTCTGCCACACCCATAAAT
 CCTTGTGTGTTAACTTCTAGCTGCCTGGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGGCAT
 CTTTGGCCAGGCTTCTGCCCCCTGCAGCTGGGACCCCTCACTTGCTGCCATGCTCTGCTCGGCT
 TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTTTAATTTGCATTTTTTTTC
 ATTTGGGGCCAAAAGTCCAGTGAAATTGTAAGCTTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYG
 HPNPGMFPSGTPGGPYGGAAPGGPYGQPPSSYGAQQPGLYGQGGAPPNVDPEAYSWFQSVDS DH
 SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTKSGRIDVYGFSALWKFIQQWKNLFQQYDR
 DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFREK
 DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-
 66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCCGCGTGGCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCATC
 TCCCAGGCCTCTTTGCCCCGAGCATCGGTGTTGTGGAGGAGAAAGTTTCCAAAACTTCGGGACC
 AACTTGCCCTCAGCTCGGACAACCTTCCTCCACTGGCCCCCTCTAACTCTGAACATCCGCAGCCCCG
 TCTGGACCCTAGGTCTAATGACTTGGCAAGGGTTCTCTGAAGCTCAGCGTGCCTCCATCAGATG
 GCTTCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGGCTGCCTGCCATG
 GATTCTGGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCGGCTGAGGACCGCCTGGGGGA
 AGCGCTGCCTGAAGAACTCTCTTACCTCTCCAGTGCTGCGGCCCTCGCTCCGGGCAGTGGCCCTT
 TGCCTGGGGAGTCTTCTCCCGATGCCACAGGCCTCTCACCTGAGGCTTCACTCCTCCACCAGGAC
 TCGGAGTCCAGACGACTGCCCCGTTCTAATTCAGTGGGAGCCGGGGGAAAAATCCTTTCCCAACG
 CCTCCCTGGTCTCTCATCCACAGGGTTCTGCCTGATCACCCCTGGGGTACCCTGAATCCAGTG
 TGTCTGGGGAGGTGGAGGCCCTGGGACTGGTTGGGGAACGAGGCCCATGCCACACCCTGAGGGA
 ATCTGGGGTATCAATAATCAACCCCCAGGTACCAGCTGGGGAAATATTAATCGGTATCCAGGAGG
 CAGCTGGGGAAATATTAATCGGTATCCAGGAGGCAGCTGGGGGAATATTAATCGGTATCCAGGAG
 GCAGCTGGGGGAATATTCATCTATACCCAGGTATCAATAACCCATTTCCTCCTGGAGTTCTCCGC
 CCTCCTGGCTCTTCTTGGAACATCCCAGCTGGCTTCCCTAATCCTCCAAGCCCTAGGTTGCAGTG
 GGGCTAGAGCACGATAGAGGGAAACCAACATTGGGAGTTAGAGTCCTGCTCCCGCCCCTTGCTG
 TGTGGGCTCAATCCAGGCCCTGTTAACATGTTTCCAGCACTATCCCCACTTTTCAGTGCCTCCCC
 TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 64

MQGRVAGSCAPLGLLLVCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGSPNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPPAGGSAVQRWPPSWGLPAMDSWPPEDPWQMMAAAEDRLGEA
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLPRSNSLGAGGKILSQRP
PWSLIHRVLPDHPWGTNLNPSVSWGGGGPGTGWGTRPMPHPEGIWGINNQPPGTSGWGNINRYPGGS
WGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPPGVLRPPGSSWNIPAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGGGTC
TGGGCTGCCCCCTTGCTCCTCTTGACCCCTCCTTGGCAGCTCACATGGAACAGGGCCGGGTATGA
CTTTGCAACTGAAGCTGAAGGAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGCTTCCTGGAA
TTGCTTGAAAAGCTCTGCCTCCTCCTCCATCTCCCTTCAGGGACCAGCGTCACCCTCCACCATGC
AAGATCTCAACACCATGTTGTCTGCAACACATGACAGCCATTGAAGCCTGTGTCCTTCTTGGCCC
GGGCTTTGGGCCGGGGATGCAGGAGGCAGGCCCCGACCCTGTCTTTCAGCAGGCCCCCACCTC
CTGAGTGGCAATAAATAAAATTCGGTATGCTG

FIGURE 66

MSGSLPLVLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSESSFLELLEKLCLLLHLPSTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCCGGGC
CAGGTGCCCCGTCGCAGGTGCCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGAAGCCCC
TTCTTCGGCGCTGCCAACCCGCCACCCAGCCCATGGCGAACCCCGGGCTGGGGCTGCTTCTGGCG
CTGGGCCTGCCGTTCTGCTGGCCCGCTGGGGCCGAGCCTGGGGGCAAATACAGACCACTTCTGC
AAATGAGAATAGCACTGTTTTGCCCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTGCTCCTGGCTGTGGGGCTG
GCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTACCGGCCCACTAGCGA
GGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
GCCTGCCCATCTAGGTCCCCCTCTCCTGCATCTGTCTCCCTTCATTGCTGTGTGACCTTGGGGAAA
GGCAGTGGCCTCTCTGGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAGAAGGTACTTCAA
AGACTCTGCCCCCTGAGGTCAAGAGAGGATGGGGCTATTCACCTTTATATATTTATATAAAATTAG
TAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

FIGURE 69

GCCAGGAATAACTAGAGAGGAACAATGGGGTTATTAGAGGTTTTGTTTTCTTCTAGTTCTGTGCCTGCTGCACCAG
 TCAAATACTTCCCTTATTAGCTGAATAATAATGGCTTTGAAGATATTGTCATTGTTATAGATCCTAGTGTGCCAGAA
 GATGAAAAAATAATTGAACAAATAGAGGATATGGTGACTACAGCTTCTACGTACCTGTTTGAAGCCACAGAAAAAGA
 TTTTTTTTCAAAAATGTATCTATATTAATTCCTGAGAATTGGAAGGAAAAATCCTCAGTACAAAAGGCCAAAACATGAA
 AACCATAAACATGCTGATGTTATAGTTGCACCACCTACACTCCCAGGTAGAGATGAACCATACCAAGCAGTTTACA
 GAATGTGGAGAGAAAGGCCAATACATTCACTTCACCCCGACCTTCTACTTGGAAAAAACAAAATGAATATGGACCA
 CCAGGCCAACTGTTTGTCCATGAGTGGGCTCACCTCCGGTGGGGAGTGTTTGATGAGTACAATGAAGATCAGCCTTTC
 TACCGTGCTAAGTCAAAAAAATCGAAGCAACAAGGTGTTCCGCAGGTATCTCTGGTAGAAATAGAGTTTATAAGTGT
 CAAGGAGGCAGCTGTCTTAGTAGAGCATGCAGAATTGATTCACAAACAACTGTATGGAAAAGATTGTCAATTCTTT
 CCTGATAAAGTACAAACAGAAAAAGCATCCATAATGTTTATGCAAAGTATTGATTCGTGTTGTTGAATTTGTAACGAA
 AAAACCCATAATCAAGAAGCTCCAAGCTACAAAACATAAAGTGCAATTTAGAAAGTACATGGGAGGTGATTAGCAAT
 TCTGAGGATTTTAAAAACACCATACCCATGGTGACACCACCTCCTCCACCTGTCTTCTCATTGCTGAAGATCAGTCAA
 AGAATTTGTGTGCTTAGTTCTTGATAAGTCTGGAAGCATGGGGGGTAAGGACCGCTAAATCGAATGAATCAAGCAGCA
 AAACATTTCTGCTGCAGACTGTTGAAAATGGATCCTGGGTGGGGATGGTTCACCTTGATAGTACTGCCACTATTGTA
 AATAAGCTAATCCAAATAAAAAGCAGTGATGAAAGAAACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGA
 ACTTCCATCTGCTCTGGAATTAAATATGCATTTCAAGGTGATTGGAGAGCTACATTCCTCACTCGATGGATCCGAAGTA
 CTGCTGCTGACTGATGGGAGGATAAAGTCTGCAAGTCTTGTATTGATGAAGTGAACAAAGTGGGGCCATTGTTTCTAT
 TTTATTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAATAGAGATGAGCAAGATAACAGGAGGAAGTCATTTTATGTT
 TCAGATGAAGCTCAGAACATGGCCTCATTGATGCTTTTGGGGCTTTACATCAGGAAATACTGATCTCTCCAGAG
 TCCCTTCAGCTCGAAAGTAAGGGATTAACTGATGAATAGTAATGCCTGGATGAACGACACTGTCAATAATGATAGTACA
 GTGGGAAGGACACGTTCTTCTCATCATGGAACAGTCTGCCTCCAGTATTCTCTCTGGGATCCAGTGGGAACA
 ATAATGGAAAAATTTACAGTGGATGCAACTTCCAAATGGCCTATCTCAGTATTCCAGGAAGTCAAAGGTGGGCACT
 TGGGCATACAACTTCAAGCCAAAGCGAACCCAGAAACATTAATATACAGTAAGTCTCGAGCAGCAAAATCTTCT
 GTGCCCTCAATCACAGTGAATGCTAAATGAATAAGGACGTAAACAGTTTCCCCAGCCCAATGATTGTTTACGCAGAA
 ATTCTACAAGGATATGTACCTGTTCTTGGAGCCAATGTGACTGCTTTCATTGAATCACAGAATGGACATACAGAAGTT
 TTGGAACTTTTGGATAATGGTGCAAGGCGCTGATTCTTCAAGAATGATGGAGTCTACTCCAGGTATTTTACAGCATAT
 ACAGAAAATGGCAGATATAGCTTAAAAGTTCCGGGCTCATGGAGGAGCAAACTGCCAGGCTAAAATTACGGCCTCCA
 CTGAATAGAGCCGCTACATACCAGGCTGGGTAGTGAACGGGGAAATGAAGCAAAACCCGCCAAGACCTGAAATTGAT
 GAGGATACTCAGACCACCTTGGAGGATTTCAAGCCGAACAGCATCCGGAGGTGCATTTGTGGTATCACAAGTCCCAAGC
 CTTCCCTTGCCTGACCAATACCCACCAAGTCAAAATCACAGACCTTGATGCCACAGTTTCATGAGGATAAGATTATTCTT
 ACATGGACAGCACCAGGAGATAATTTTGTATGTTGGAAAAGTTCAACGTTATATCATAAGAATAAGTGCAAGTATTCTT
 GATCTAAGAGACAGTTTGTATGATGCTCTTCAAGTAAATACTACTGATCTGTACCAAAGGAGGCCAACTCCAAGGAA
 AGCTTTGCATTTAAACAGAAAAATATCTCAGAAGAAAATGCAACCCACATATTTATTGCCATTAAAAGTATAGATAAA
 AGCAATTTGACATCAAAAGTATCCAACATTGCACAAGTAACTTTGTTTATCCCTCAAGCAAACTCCTGATGACATTGAT
 CCTACACCTACTCCTACTCCTACTCCTGATAAAAGTCATAATTCTGGAGTTAATTTTCTACGCTGGTATTG
 TCTGTGATTGGGTCTGTTGTAATTGTTAACTTTATTTAAGTACCACCATTTGAACCTTAACGAAGAAAAAATCTTC
 AAGTAGACCTAGAAGAGAGTTTTAAAAACAAAAAATGTAAGTAAAGGATATTCTGAATCTTAAATTCATCCCAT
 GTGTGATCATAAACTCATAAAATAATTTAAGATGTGGAAGGATACTTTGATTAATAAAAAACACTCATGGATA
 TGTAAAACTGTCAAGATTAATAATTAATAGTTTCATTTATTTGTTATTTTATTGTAAGAAATAGTGATGAACAAAG
 ATCCTTTTTCATACTGATACCTGGTTGTATATTATTGATGCAACAGTTTTCTGAAATGATATTTCAAATTCATCAA
 GAAATTAATCATCTATCTGAGTAGTCAAAATACAAGTAAAGGAGAGCAAAATAAACACATTTGGAAAAAATAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 70

MGLFRGFVFLVLCLLHQSNTSFIKLNNNGFEDIVIVIDPSVPEDEKIEQIEDMVTASTYLFE
 ATEKRFFFKNVSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTECGEKGEY
 IHFTPDLLLQKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRCISAGISGRN
 RVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDQVQTEKASIMFMQSIDSVVEFCNEKTHNQEAP
 SLQNIKCNRSTWEVISNSEDKNITPMVTPPPPVFSLKISQRIVCLVLDKSGSMGGKDRNLNR
 MNQAAKHFLLOTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSICSGIK
 YAFQVIGELHSQLDGSEVLLLTGDNATSSCIDEVKQSGAIVHFIALGRAADEAVIEMSKITGG
 SHFYVSDEAQNNGLIDAFGALTSGNTDLSQSLQLESKGLTLNSNAWMNDTVIIDSTVGKDTFFL
 ITWNSLPPSISLWDPSTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKANPETLTITVTSR
 AANSSVPPITVNAKMNDVNSFPSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA
 GADSFKNDGVYSRYFTAYTENGRYSLKVRHGGANTARLKLRLPLNRAAYIPGWVVNGEIEANPP
 RPEIDEDTQTTLEDFSRASGGAFVVSQVPSLPLPDQYPPSQITDLDTVHEDKIIILTWTAPGDN
 FDVGKVQRYIIRISASILDRLDSFDDALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAI
 KSIDKSNLTSKVSNIAQVTLFIPQANPDDIDPTPTPTPTPTPKSHNSGVNISTLVLSVIGSVVI
 VNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
 628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGGTGGAAACCCCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACGTCCTCCCGGGCAGGGGTGA
CAACAGGTGTCACTTTTTGATCTCGTGTGTGGCTGCCTTCTATTTCAAGGAAAGACGCCAAGGTAATTTGACCCA
GAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTCTTGAACCCCAAGTTATGCCAGGATTACTAGAGAGTGTCA
ACTCAACCAGCAAGCGGCTCCTTCGGCTTAACTTGTGGTTGGAGGAGAGAACCCTTGTGGGGCTGCGTTCTCTTAGCA
GTGCTCAGAAGTGACTTGCCTGAGGGTGGACCAGAAGAAAGGAAAGGTCCCCCTTGTGCTGTTGGCTGCACATCAGGAA
GGCTGTGATGGGAATGAAGGTGAAACTTGGAGATTTCACTTCAGTCATTGCTTCTGCCTGCAAGATCATCTTTAAA
AGTAGAGAAGCTGCTCTGTGTGGTGGTTAACTCCAAGAGGCAGAACTCGTTCTAGAAGGAAATGGATGCAAGCAGCTC
CGGGGGCCCCAAACGCATGCTTCTGTGGTCTAGCCCAGGGAAGCCCTTCCGTGGGGGGCCCCGGCTTTGAGGGATGCC
ACCGGTTCTGGACGCATGGCTGATTCTGAATGATGATGGTTTCCGCGGGGGCTGCTTGCCTGGATTTCCTGGGTGGTG
GTTTTGCTGGTGCTCCTCTGCTGTGCTATCTCTGTCTGTACATGTTGGCTGCACCCAAAAGGTGACGAGGAGCAG
CTGGCACTGCCAGGCCAACAGCCCCACGGGGAAGGAGGGGTACAGGCCCTCCTTCAGGAGTGGGAGGAGCAGCAC
CGCAACTACGTGAGCAGCTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGTGCAGGAGAGGAGTGCAGCTCAGG
AATGGGCAGTACCAAGCCAGCGATGCTGCTGGCTGGGTCTGGACAGGAGCCCCCAGAGAAAAACCCAGGCCGACCTC
CTGGCCTTCTGCACTCGCAGGTGGACAAGGCAGAGGTGAATGCTGGCTCAAGCTGGCCACAGAGTATGCAGCAGTG
CCTTTCGATAGCTTTACTCTACAGAAGGTGTACCAGCTGGAGACTGGCCTTACCCGCCACCCGAGGAGAAGCCTGTG
AGGAAGGACAAGCGGGATGAGTTGGTGGAAAGCCATTGAATCAGCCTTGGAGACCTGAACAATCTGCAGAGAACAGC
CCCAATCACCCTCCTTACAGCGGCTCTGATTTATAGAAGGGATCTACCGAACAGAAAGGACAAAGGCACATTGTAT
GAGCTCACCTTCAAAGGGGACCACAAACAGAACTCAAACGGCTCATCTTATTCGACCATTCAGCCCCATCATGAAA
GTGAAAAATGAAAAGCTCAACATGGCCAAACAGCTTATCAATGTTATCGTGCCTTAGCAAAAAGGGTGGACAAGTTC
CGGCAGTTCATGCAGAAATTCAGGGAGATGTGCATTGAGCAGGATGGGAGAGTCCATCTCACTGTTGTTTACTTTGGG
AAAGAAGAAATAAATGAAGTCAAAGGAATACTTGAACACACTTCCAAGCTGCCAACTTCAGGAACCTTACCTTCATC
CAGCTGAATGGAGAATTTCTCGGGGAAAGGACTTCAATGTTGGAGCCCGCTTCTGGAAGGGAAGCAAGCTCCTTCTC
TTTTTCTGTGATGTGGACATCTACTTCACTCTGAATTCCTCAATACGTGTAGGCTGAATACACAGCCAGGGAAGAAG
GTATTTTATCCAGTTCTTTTCACTCAGTACAATCCTGGCATAATATACGGCCACCATGATGCGATCCCTCCCTTGGAA
CAGCAGCTGGTCATAAAGAAGGAACTGGATTTTGGAGAGACTTGGATTGGGATGACGTGTCAATCTCGGTGAGAC
TTCATCAATATAGGTGGTTTGTATCTGGACATCAAAGGCTGGGGCGGAGAGGATGTGCACCTTTATCGCAAGTATCTC
CACAGCAACCTCATAGTGGTACGGACGCTGTGGAGGACTCTTCCACCTCTGGCATGAGAAGCGGTGCATGGACGAG
CTGACCCCGAGCAGTACAAGATGTGCATGCAGTCCAAGGCCATGAACGAGGCATCCACGGCCAGCTGGGCATGCTG
GTGTTTCAGGCACGAGATAGAGGCTCACCTTCGCAACAGAAACAGAAAGCAAGTAGCAAAAAACATGAACCTCCAGA
GAAGGATTGTGGGAGACACTTTTCTTCTCTTTTGAATTAAGTGGCTGCAACAGAGAAAAAGACTTCCATAAA
GGACGACAAAAGAAATGGACTGATGGGTGAGAGATGAGAAAGCCTCCGATTTCTCTCTGTTGGGCTTTTACAAACAGA
AATCAAAATCTCCGCTTTCCTGCAAAAGTAACCCAGTTGCACCTGTGAAGTGTCTGACAAAGGCAGAAATGCTTGTG
AGATTATAAGCCTAATGGTGTGGAGGTTTGTATGGTGTATACAATACACTGAGACCTGTTGTTTTGTGTCTCATTTGA
AATATTCATGATTTAAGAGCAGTTTGTAAAAAATTCATTAGCATGAAAGCAAGCATATTTCTCCTCATATGAATGA
GCCTATCAGCAGGGCTCTAGTTTCTAGGAATGCTAAAATATCAGAAGGCAGGAGAGGATAGGCTTATTATGATACT
AGTGAGTACATTAAGTAAAAATAAATGGACCAGAAAAAGAAAGAACATAAATATCGTGTCAATTTTCCCCAAGAT
TAACCAAAATAATCTGCTTATCTTTTGGTTGTCTTTTAACTGTCTCCGTTTTTTCTTTTATTTAAAAATGCACT
TTTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATTACCACTTTGCAAGCCTTACAAGAGAGCAAGTTGGCCTAC
ATTTTTATATTTTTTAAGAAGATACTTTGAGATGCATTATGAGAACTTTCAGTTCAAAGCATCAAATGTATGCCATAT
CCAAGGACATGCCAAATGCTGATTCTGTGAGGCACTGAATGTGAGGCATTGAGACATAGGGAAGGAATGGTTTGTACT
AATACAGACGTACAGATACTTCTCTGAAGAGTATTTTGAAGAGGAGCAACTGAACACTGGAGGAAAAAGAAATGAC
ACTTCTGCTTTACAGAAAAGGAACTCATTGAGACTGGTGATATCGTGATGATCTTAAAGTCAGAAACCAATTTT
CTCCTCAGAAAGTAGGGACCGCTTCTTACCTGTTTAAATAAACCAAGTATACCGTGTGAACCAACAATCTCTTTTC
AAAAAGGGTGTCTCCTGGCTTCTGGCTTCCATAAGAAGAAATGGAGAAAAATATATATATATATATATATTTGT
GAAAGATCAATCCATCTGCCAGAACTAGTGGGATGGAAGTTTTTGTACATGTTATCCACCCAGGCCAGGTGGAAG
TAACTGAATATTTTTTAAATTAAGCAGTTCTACTCAATCACCAGATGCTTCTGAAATTGCAATTTTATACCATTT
CAAACTATTTTTTAAAAATAAATACAGTTAATACATAGAGTGGTTTCTTCAATCATGTGAAAAATTATAGCCAGCACCAG
ATGCATGAGCTAATATCTCTTGAAGTCTGCTTCTGTTTGTCTCACAGTAAACTCATTTTAAAGCTTCAAGAAC
ATTCAAGCTGTTGGTGTGTTAAAAATGCATTGTATTGATTGTTACTGGTAGTTTATGAAATTTAATTAACACAGG
CCATGAATGGAAGGTGGTATTGCACAGCTAATAAATATGATTGTGGATATGAA

FIGURE 72

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEW
 EQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK
 AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVKDKRDELVEAIESALETNNPA
 ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVKNELNMAN
 TLINIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVKGILENTSKAANFRNF
 TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVFYPVLFSQY
 NPGIIYGHHDVAPPLEQQQLVIKKETGFWRDFGFGMTCQYRSDFINIGGFDDIKGWGGEDVHLYR
 KYLHNSLIVVRTFVRGLFHLWHEKRCMDELTPQYKMCMSKAMNEASHGQLGMLVFRHEIEAHL
 RKQKQKTSSKKT

Important features:**Signal peptide:**

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCTGTGTC
 TCTGGTGGTTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGAATGTCCTACA
 ATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTATCACAA
 ATTGGCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGTCCTGACAGCTCC
 AGAGAAGTGAAGAGAAATCCAGAAGACCTTCTGTTTCCATGCAACAAATATACTCCAATCTGA
 AGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAACCAC
 ACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTGGAGTCCTTCGTCCC
 AGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGATCAATCAT
 CAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGTGTTCCTTTT
 TCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACACCCAGCAAATTT
 GATTTTGATTTATGGAAATGAATTTGACAAAAGATTCTTTGTGCCTGCTGAAAAAATCGTGATTA
 ACTTTATCACCCCTCAATATCTCGGATGATTCTAAAAATTTCTCATCAGGATATGAGTTTACTGGGA
 AAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGA
 GGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGATGGAAATTTTTTGTGACTCTGAAG
 AAAACACGGAAGGTACTTCTCTCACCCAGCAAGAGTCCCTCAGCAGAACAAATACCCCGGATAAA
 ACAGTCATTGAATATGAATATGATGTCAGAACCACTGACATTTGTGCGGGGCCTGAAGAGCAGGA
 GCTCAGTTTGCAGGAGGAGGTGTCCACACAAGGAACATTATTGGAGTCGAGGCAGCGTTGGCAG
 TCTTGGGCCCGCAAACGTTACAGTACTCATAACCCCTCAGCTCCAAGACTTAGACCCCTGGCG
 CAGGAGCACACAGACTCGGAGGAGGGCCGGAGGAAGAGCCATCGACGACCCTGGTCGACTGGGA
 TCCCCAAACTGGCAGGCTGTGTATTCTTCGCTGTCCAGCTTCGACCAGGATTCAGAGGGCTGCG
 AGCCTTCTGAGGGGGATGGGCTCGGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGCT
 CCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGGTTATA
 TGTGCAGATGGAAAACTGATGCCAACACTTCCTTTTGCCTTTTGTTCCTGTGCAAACAAGTGAG
 TCACCCCTTTGATCCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTTTCCAGTTTGTGAGTGT
 CTGTGAGAATTACTTATTTCTTTCTCTATTCTCATAGCACGTGTGTGATTGGTTCATGCATGTA
 GGTCTCTTAACAATGATGGTGGGCTCTGGAGTCCAGGGGCTGGCCGGTTGTTCTATGCAGAGAA
 AGCAGTCAATAAATGTTTGGCAGACTGGGTGCAGAATTTATTCAGGTGGGTGT

FIGURE 76

MSYNGLHQRVFKELKLLTLCSSISQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIY
 SNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRAQPSÉKQCARTLK
 DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNFDRFFVPAEK
 IVINFITLNISDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFC
 DSEENTEGTSLTQQESLSRTIPDKTVIEYDYVRTTDICAGPEEQELSLQEEVSTQGTLLESQA
 ALAVLGPQTLQYSYTPQLQDLPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRLCIPSLSSFDQDS
 EGCEPSEGDLGEEGLLSRLYEPPAPDRPPGENETYLMQFMEEWGLYVQMEN

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTGCCAGGTCTGGCATCCTGCACTTGCTGCCCTCTGACAC
 CTGGGAAGATGGCCGGCCCGTGACCTTCACCTTCTCTGTGGTTTGCTGGCAGCCACCTTGATC
 CAAGCCACCCTCAGTCCCACTGCAGTTCTCATCCTCGGCCAAAAGTCATCAAGAAAAGCTGAC
 ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC
 GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACCGTCCTGAAGCACATC
 ATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAAGCCCTCGGCCAATGACCA
 GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA
 TCGTGGAGTTCCACATGACGACTGAGGCCAAGCCACCATCCGCATGGACACCAGTGCAAGTGGC
 CCCACCCGCTGGTCCCTCAGTGACTGTGCCACCAGCCATGGGAGCCTGCGCATCCAAGTGTGTA
 TAAGCTCTCCTTCTGGTGAACGCCTTAGCTAAGCAGGTGATGAACCTCCTAGTGCCATCCCTGC
 CCAATCTAGTGAAAAACCAGCTGTGTCCCGTGATCGAGGCTTCCCTTCAATGGCATGTATGCAGAC
 CTCTGCAGCTGGTGAAGGTGCCATTTCCTCAGCATTGACCGTCTGGAGTTTGACCTTCTGTA
 TCCTGCCATCAAGGTGACACCATTGAGCTCTACCTGGGGGCCAAGTTGTTGGACTCACAGGGAA
 AGGTGACCAAGTGTTCAATAACTCTGCAGCTTCCCTGACAATGCCACCCTGGACAACATCCCG
 TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGA
 ATTCATGGTCTGTGGACTCTGTGCTTCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGC
 TGATCAATGAAAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGAC
 ACTCCCGAGTTTTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGTT
 TCCCTCCAGTGAAGCCCTCCGCCCTTTGTTACCCCTGGGCATCGAAGCCAGCTCGGAAGCTCAGT
 TTACACCAAAGGTGACCAACTTATACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTG
 ATGAACTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA
 CTCCATCCTGCTGCCGAACCAGAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGTGAAGG
 CCTTGGGATTCGAGGCAGCTGAGTCCTCACTGACCAAGGATGCCCTTGTGCTTACTCCAGCCTCC
 TTGTGGAAACCCAGCTCTCTGTCTCCCAGTGAAGACTTGGATGGCAGCCATCAGGGAAGGCTGG
 GTCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAATCAATAAACAATTG
 CCTGTGAAAAA

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIEKLTQELKDHNATSILQQPLLSAMREK
 PAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE
 FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLENL
 VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDITQLYLGAQLLDSQKVT
 KWFNNSAASLTMPITLDNIPFSLIVSQDVVKAAVAVALSPEEFMVLLDSVLPESAHLKSSIGLIN
 EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSSEALRPLFTLGIEASSEAQFYT
 KGDQLILNLNNISSDRIQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSQVPSLVKALG
 FEAAESSLTKDALVLTTPASLWKPSSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGGCCTCT
CTTGGCCTCCAACCTTGTTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACTGGTTGCCAT
GCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCCAGCATTGTGACAGCAGTTGGCTTCT
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC
ACCCCTTCTGGGCCTGCCCCTGACATCCAGGCTGCCAGGCCATGATGGTGACATCCAGTGCAAT
CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCACAGTCTTCTGCCAGGAATCCC
GAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTTCATCCTTGGAGGCCTCCTGGGATTC
ATTCTGTGCTGGAATCTTCATGGGATCCTACGGGACTTCTACTCACCCTGGTGCCTGACAG
CATGAAATTTGAGATTGGAGAGGCTCTTACTTGGGCATTATTTCTTCCCTGTTCTCCCTGATAG
CTGGAATCATCCTCTGCTTTTCCTGCTCATCCCAGAGAAATCGCTCCAATACTACGATGCCTAC
CAAGCCCAACCTCTTGCCACAAGGAGCTCTCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGA
GTTCAATTCCTACAGCCTGACAGGGTATGTGTGAAGAACCAGGGGCCAGAGCTGGGGGGTGGCTG
GGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCCACAGGTGAGGGACACTACCACTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGG
GCTAGTGTAACAGCATGCAGGTTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTCTGTTTTCC
TCACCTTGCTGCTCCCTGCCCTAAGTCCCCAACCTCAACTTGAAACCCCATTCCTTAAGCCA
GGACTCAGAGGATCCCTTTGCCCTCTGGTTTACCTGGGACTCCATCCCCAAACCCACTAATCACA
TCCCCTGACTGACCCTCTGTGATCAAAGACCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCAT
GCTGGGGATGGGAAGGAGAAGCAGTGGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTC
CCTCCAAAGAACTGATTGGCCCTGGAACCTCCATCCCCTCTTGTATGACTCCACAGTGTCCA
GACTAATTTGTGCATGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAATA

FIGURE 8o

MASLGLQLVGYILGLLGLLGTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAAQAMMTSSAIISSLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPVAWNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVKSEFNSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCTTTC
 CCCGCGTTCTCTTCCACCTTTCTCTTCTTCCCACCTTAGACCTCCCTTCTGCCCCTCCTTTCCT
 GCCCACCGCTGCTTCCTGGCCCTTCTCCGACCCCGCTCTAGCAGCAGACCTCCTGGGGTCTGTGG
 GTTGATCTGTGGCCCTGTGCCTCCGTGCCTTTTCGTCTCCCTTCTCCCGACTCCGCTCCCGG
 ACCAGCGGCCTGACCCTGGGGAAGGATGTTCCCGAGGTGAGGGTCTCTCCTCCTTGTCTGGGA
 CTCGCGCTGCTCTGGTTCCCCCTGGACTCCCACGCTCGAGCCCGCCAGACATGTTCTGCCTTTT
 CCATGGGAAGAGATACTCCCCGGCGAGAGCTGGCACCCCTACTTGGAGCCACAAGGCCTGATGT
 ACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGTTGTTACCGCTCCACTGTCCGCCT
 GTCCACTGCCCCCAGCCTGTGACGGAGCCACAGCAATGCTGTCCCAAGTGTGTGGAACCTCACAC
 TCCCTCTGGACTCCGGGCCCCACCAAAGTCTGCCAGCACAACGGGACCATGTACCAACACGGAG
 AGATCTTCAGTGCCCATGAGCTGTTCCCTCCCGCTGCCCAACCAGTGTGTCTCTGCAGCTGC
 ACAGAGGGCCAGATCTACTGCGGCCTCACAACCTGCCCCGAACCAGGCTGCCAGCACCCCTCCC
 ACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGCAAGTGAGCAATCGGATGAAGAGGACA
 GTGTGCAGTCGCTCCATGGGGTGAGACATCCTCAGGATCCATGTTCCAGTGATGCTGGGAGAAAG
 AGAGGCCCGGGCACCCAGCCCCACTGGCCTCAGCGCCCTCTGAGCTTCATCCCTCGCCACTT
 CAGACCCAAAGGAGCAGGCAGCACAACGTCAAGATCGTCCTGAAGGAGAAACATAAGAAAGCCT
 GTGTGCATGGCGGAAGACGTACTCCACGGGGAGGTGTGGCACCCGGCCTTCCGTGCCTTCGGC
 CCCTTGCCCTGCATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCC
 CACCGAGTACCCCTGCCGTACCCCGAGAAAGTGGCTGGGAAGTGCTGCAAGATTTGCCAGAGG
 ACAAAGCAGACCCTGGCCACAGTGAGATCAGTTCTACCAGGTGTCCCAAGGCACCGGGCCGGGTC
 CTCGTCCACACATCGGTATCCCCAAGCCCAGACAACCTGCGTCGCTTTGCCCTGGAACACGAGGC
 CTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAAGATGAGGAACTGAGGCTCAGAGAG
 GTGAAGTACCTGGCCCAAGGCCACACAGCCAGAATCTTCCACTTGACTCAGATCAAGAAAGTCAG
 GAAGCAAGACTTCCAGAAAGAGGCACAGCACTTCCGACTGCTCGCTGGCCCCCACGAAGGTCCT
 GGAACGTCTTCTAGCCCAGACCCTGGAGCTGAAGGTCACGGCCAGTCCAGACAAAGTGACCAAG
 ACATAACAAAGACCTAACAGTTGCAGATATGAGCTGTATAATTGTTATTATATATTAATAAA
 TAAGAAGTTGCATTACCCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARAREDMFCLFHGKRYPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPPVHCPQPVTEPQQCCPKCEPHTPSGLRAPPKSCQHNGTMYQHGEI FSAHELF
PSRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPPTPAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKHKKACVHGGKTY
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTTEPCRHPKVAGKCKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH
SQNLPLDSDQESQEARLPERGTALPTARWPPRRSLERLPSPDPGAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTGTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGCCCTCCGCTCAGCGAGAGCCTCTCC
GTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTCTCTCTCTAATCCATCCGTACCTCTCCTGTCA
TCCGTTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTCTCATGCTCAGTTTGTTCTGAGTC
TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTTGGGCCAGACAAGCCTGTCCAGGCCTTGTTGGGGGAG
GACGCAGCATTCTCCTGTTTCTGTCTCCTAAGACCAATGCAGAGGCCATGGAAGTGCGGTTCTTCAGGGG
CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTTATGCAGATGCCACAGTATC
AAGGCAGGACAAAACGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCTGAGGCTGGAAAACATTACT
GTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCCAGTCTTACTACCAGAAGGCCATCTGGGAGCT
ACAGGTGTGAGCACTGGGCTCAGTTCTCTCATTTCCATCACGGGATATGTTGATAGAGACATCCAGCTAC
TCTGTGAGTCTCGGGCTGGTTCCCCCGGCCACAGCGAAGTGGAAGGTCCACAAGGACAGGATTTGTCC
ACAGACTCCAGGACAAACAGAGACATGCATGGCTGTTTGATGTGGAGATCTCTGTACCGTCCAAGAGAA
CGCCGGGAGCATATCTGTTCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG
GAGATACCTTTTTCGAGCCTATATCGTGGCACCTGGCTACCAAAGTACTGGGAATACTCTGCTGTGCCTA
TTTTTTGGCATTGTTGGACTGAAGATTTCTTCTCCAAATCCAGTGGAATCCAGGCGGAACTGGACTG
GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCGGAAACACGCAGTGAGGTGACTCTGGATCCAG
AGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAAACGTAAACCCATAGAAAAGCTCCCCAGGAGGTG
CCTCACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCTTCTCAGAGTTTCCAAGCAGGGAAACATTA
CTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGCGTGGGAGTGTGCCGGGATGATGTGGACAGGAGGA
AGGAGTACGTGACTTTGTCTCCCGATCATGGGTACTGGGTCTCAGACTGAATGGAGAACATTTGTATTTC
ACATTAAATCCCCGTTTATCAGCGTCTTCCCCAGGACCCACCTACAAAAATAGGGGTCTTCTGGACTA
TGAGTGTGGGACCATCTCCTTCTTCAACATAAATGACCAGTCCCTTATTTATACCCTGACATGTGGTTTG
AAGGCTTATTGAGGCCCTACATTGAGTATCCGTCTATAATGAGCAAAATGGAATCCCATAGTCATCTGC
CCAGTCACCCAGGAATCAGAGAAAGAGGCCTCTTGGCAAAGGGCCTCTGCAATCCAGAGACAAGCAACAG
TGAGTCTCTCACAGGCAACACGCCCTTCTCCCCAGGGGTGAAATGTAGGATGAAATCACATCCCACAT
TCTTCTTTAGGGATATTAAGGTCTCTCTCCAGATCCAAAGTCCCGCAGCAGCCGGCCAAGGTGGCTTCCA
GATGAAGGGGACTGGCCTGTCCACATGGGAGTCAGGTGTCATGGCTGCCCTGAGCTGGGAGGGAAGAAGG
CTGACATTACATTTAGTTTGCTCTCACTCCATCTGGCTAAGTGATCTTGAAATACCACCTCTCAGGTGAAG
AACCGTCAGGAATCCCATCTCACAGGCTGTGGTGTAGATTAAGTAGACAAGGAATGTGAATAATGCTTAG
ATCTTATTGATGACAGAGTGTATCCTAATGGTTGTTCATTATATTACACTTTCAGTAAAAAAA

FIGURE 84

MALMLSLVLSLLKLGSGQWQVFGPDKPVQALVGEDAAFSCFLSPKTNAEAMEVRFFRGQFSSVVH
 LYRDGKDQPFMQMPQYQGRTKLVKDSIAEGRISLRLENITVLDAAGLYGCRISSQSYQKAIWELQ
 VSALGSVPLISITGYVDRDIQLLCQSSGWFPRTAKWKGPQGDLSTDSRTNRDMHGLFDVEISL
 TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
 FQWKIQAELDWRRKHGQAEIRDARKHAVEVTLPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF
 TRKSVVASQSFQAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLSPDHGYWVLRNLNGEHLFT
 LNPRFISVFPRTPTTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPHYIEYPSYNEQNGT
 PIVICPVTQESEKEASWQRASAIPESTSNESSSQATTPFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCAGACATGCTGCTGCTGCTGCTGCCCCT
GCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTTCCGTGA
CGGTGCAGGAAGGCCTGTGTGTCCATGTGCCCTGCTCCTTCTCCTACCCCTCGCATGGCTGGATT
TACCCTGGCCCAGTAGTTCATGGCTACTGGTTCCGGAAGGGGCCAATACAGACCAGGATGCTCC
AGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTCCACCTCCTTG
GGGACCCACATACCAAGAATTGCACCCTGAGCATCAGAGATGCCAGAAGAAGTGATGCGGGGAGA
TACTTCTTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACCGGCTCTCTGTGAA
TGTGACAGCCTTGACCCACAGGCCAACATCCTCATCCAGGCACCCTGGAGTCCGGCTGCCCCC
AGAATCTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAGGGGACACCCCTATGATCTCCTGGATA
GGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCCGCTCCTCGGTGCTCACCTCATCCACACA
GCCCCAGGACCATGGCACCAGCCTCACCTGTGAGGTGACCTTCCCTGGGGCCAGCGTGACCACGA
ACAAGACCGTCCATCTCAACGTGTCTACCCGCCTCAGAACTTGACCATGACTGTCTTCCAAGGA
GACGGCACAGTATCCACAGTCTTGGGAAATGGCTCATCTCTGTCACTCCCAGAGGGCCAGTCTCT
GCGCCTGGTCTGTGCAGTTGATGCAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA
GAGGCCTGACCCTGTGCCCTCACAGCCCTCAAACCCGGGGGTGCTGGAGCTGCCTTGGGTGCAC
CTGAGGGATGCAGCTGAATTCACCTGCAGAGCTCAGAACCCCTCTCGGCTCTCAGCAGGTCTACCT
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTTCGGGGGAGCTGGAG
CCACAGCCCTGGTCTTCTGTCTTCTGCGTCATCTTCGTTGTAGTGAGGTCTGCAGGAAGAAA
TCGGCAAGGCCAGCAGCGGGCGTGGGAGATACGGGCATAGAGGATGCAAACGCTGTGAGGGGTTT
AGCCTCTCAGGGGCCCCTGACTGAACCTTGGGCAGAAGACAGTCCCCCAGACCAGCCTCCCCAG
CTTCTGCCCCGCTCCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGTG
AAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG
ATGAGAAACTGCAGAGACTCACCTGATTGAGGGATCACAGCCCCCTCCAGGCAAGGGAGAAGTCA
GAGGCTGATTCTTGTAGAATTAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTATG
TGCAGAGTGAAAAGCACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCCC
TCCCTTTTATTTTTTAACTAAAAGACAGACAAATTCCTA

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTGTAC
 CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCCTGCTGTTTC
 TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT
 TCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCCTAGTGCAATTTGA
 TGGCCTGTATTTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCTGTGACATGACCTCTG
 GGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGGAAGTGCACGGTG
 GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGGACGGCAACTGGGC
 CAACTACAACACCTTTGGATCTGCAGAGGGCGGCCACGAGCGATGACTACAAGAACCCTGGCTACT
 ACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCCAATAAGTCCCCCATGCAGACTGG
 AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTCCAGACACTGGGACATAATCT
 GTTTGGCATCTACCAGAAATATCCAGTGAAATATGGAGAAGGAAAGTGTGGACTGACAACGGCC
 CGGTGATCCCTGTGGTCTATGATTTTGGCGACGCCAGAAAACAGCATCTTATTACTACCCCTAT
 GGCCAGCGGAATTCAGTGCAGGATTTGTTTCAGTTCAGGGTATTTAATAACGAGAGAGCAGCCAA
 CGCCTTGTGTGCTGGAATGAGGGTCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG
 GATACTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTTCTGGTTTTGATTGGAGTGGATAT
 GGAACATCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
 TTGAAGAGTTTGTGGGAGGGAACCCAGACCTCTCCTCCCAACCATGAGATCCCAAGGATGGAGAA
 CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGA
 AAAAAA

FIGURE 88

MNQLSFLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
 YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNFTGSAEAAAT
 SDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
 EGKCWTDNGFPVIPVVYDFGDAQKTASYSPYQGREFTAGFVQFRVFNNERAANALCAGMRVTGCN
 TEHHCIGGGGYFPEASPPQCGDFSGFDWSGYGTHVGYSSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTTGTCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCAGCCTCAGAGAC
CGCCGCCCTTGTCCCGAGGGCCATGGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTCCTGACG
CTCCTGGCGCATCTGGTGGTCGTCATCACCTTATTCTGGTCCCGGGACAGCAACATACAGGCCTG
CCTGCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTCAGCTGGTGGCCGCGCTCT
CTGTCACCCTGGGCCTCTTTGCAGTGGAGCTGGCCGGTTTCTCTCAGGAGTCTCCATGTTCAAC
AGCACCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGTGGCCCTGTCCTTCTTCAT
ATTTCGAGCGTTGGGAGTGCACTACGTATTGGTACATTTTGTCTTCTGCAGTGCCCTTCCAGCTG
TCACTGAAATGGCTTTATTCGTCACCGTCTTTGGGCTGAAAAGAAACCTTCTTGATTACCTTCA
TGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGGCCGCTTCGTATTCCTGGAAGAAGGAAG
GCATAGGCTTCGGTTTTCCCTCGGAAACTGCTTCTGCTGGAGGATATGTGTTGGAATAATTACG
TCTTGAGTCTGGGATTATCCGCATTGTATTTAGTGCCTTGTAAATAAAATATGTTTTGTAGTAACA
TTAAGACTTATATACAGTTTtaggggacaattaaaaaaaaaaaa

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACPLTFTPEEYDKQDIQLVAALSVTLGLFA
 VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCALPAVTEMALFV
 TVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

FIGURE 91

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCAACTGAC
 TCGCTGCTGCTTCGTGTTCTCTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATG
 GTCTCCCGGCTCAGAGGACCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCCCGGGTGCCT
 CGGAAGCGGGGCCACATCTACCTAAGTCCCGCCCCATGGCCAATTCCACTCTCTAGGGCTGCT
 GGCCCCGCCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACCGCCGAACCACAGCCCCC
 CACCCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTACTCCAACATCAAGACGGTG
 GCCCTGAACCTGCTCGTCACAGGGAAGATTGTGGACCATGGCAATGGGACCTTCAGCGTCCACTT
 CCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCCTCGTGCCCCCAGTAAAGCTGTAG
 AGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAGGCCTCCAAAATCTTCAACTGCCGGATG
 GAGTGGGAGAAGGTAGAACGGGGCCCGCCGACCTCGCTTTCACCCACGACCCAGCCAAGATCTG
 CTCCCGAGACCACGCTCAGAGCTCAGCCACCTGGAGCTGCTCCAGCCCTTCAAAGTCGTCTGTG
 TCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAACCTAC
 CATAGTGATACCCCCTACTACCCATCTGGGTGACCCGGGGCAGGCCACAGAGGCCAGGCCAGGGC
 TGGAAGGACAGGCCTGCCCATGCAGGAGACCATCTGGACACCGGGCAGGGAAGGGGTTGGGCCTC
 AGGCAGGGAGGGGGTGGAGACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAG
 AGAAAGGGTCCCAAGTGCTGGTCCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGG
 AGGAGGAGTGGGCTCTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGCCACAGAGAGATGCTGGG
 TCCCCGAGGCCTGTGGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGC
 CTTTGGTTCTTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGGAGATTTTCATCAGTGTGGACA
 GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCCTAGGAGCCAGTCAGCAGGGTGGGGTGGG
 GCCAGAGGAGCTCTCCAGCCCTGCCTAGTGGGCGCCCTGAGCCCCCTTGTCGTGTGCTGAGCATGG
 CATGAGGCTGAAGTGGAACCCCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC
 CAGGCCACCCCTTTCCAAAATTCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATTGCTGATG
 GCACACCCATCCTTAAGCTAAGACAGGACGATTGTGGTCTCCACACTAAGGCCACAGCCCATC
 CGCGTGCTGTGTGCTCCTTCCACCCCCAACCCCTGCTGGCTCCTCTGGGAGCATCCATGTCCCG
 GAGAGGGGTCCCTCAACAGTCAGCCTCACCTGTGACACGGGGTTCTCCCGGATCTGGATGGCGC
 CGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGGCCGGCCGAGAGCATGTGCTGGATCTGTTT
 TGTGTGTCTGTCTGTGGGTGGGGGAGGGGAGGGAAGTCTGTGAAACCGCTGATTGCTGACTTT
 TGTGTGAAGAATCGTGTTCTTGGAGCAGGAAATAAAGCTTGCCCCGGGGCA

FIGURE 92

MQLTRCCFVFLVQGSLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGNGTF
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMEWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWSCSQPFKVVCVYIAFYSTDYRLVQKVC PDYNYHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCTCGCTCGCC
CTTTATGTCTTCACCATCGCCATCGAGCCGTTGCGTATCATCTTCCTCATCGCCGGAGCTTTCTT
CTGGTTGGTGTCTCTACTGATTTTCGTCCCTTGTTTGGTTCATGGCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTCTGTCTATATCCAAGAA
ATGTTCCGATTTGCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTTGAAGAGTATAAACCC
AGGTGAGACAGCACCTCTATGCGACTGCTGGCCTATGTTTCTGGCTGGGCTTTGGAATCATGA
GTGGAGTATTTTCCTTTGTGAATACCCTATCTGACTCCTTGGGGCCAGGCACAGTGGGCATTTCAT
GGAGATTCTCCTCAATTCTCCTTTATTCAGCTTTCATGACGCTGGTCATTATCTTGCTGCATGT
ATTCTGGGGCATTGTATTTTTTGATGGCTGTGAGAAGAAAAAGTGGGGCATCCTCCTTATCGTTC
TCCTGACCCACCTGCTGGTGTGAGCCAGACCTTCATAAGTTCTTATTATGGAATAAACCTGGCG
TCAGCATTTATAATCCTGGTGCTCATGGGCACCTGGGCATTCTTAGCTGCGGGAGGCAGCTGCCG
AAGCCTGAAACTCTGCCTGCTCTGCCAAGACAAGAACTTTCTTCTTTACAACCAGCGCTCCAGAT
AACCTCAGGGAACCAGCACTTCCCAAACCGCAGACTACATCTTTAGAGGAAGCACAACTGTGCCT
TTTTCTGAAAATCCCTTTTCTGGTGGAATTGAGAAAGAAATAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAGFPALALYVFTTIAIEPLRIIFLIAGAFFWLVSLLISLVWFMARVIIDNKDG
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV
FSFVNTLSDSLPGTVGIHGDSPOFFLYSAFMTLVIILLHVFWGIVFFDGCEKKKWGILLIVLLT
HLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSCLKCLLCQDKNFLLYNQSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTTCACCAGAGTAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTTGCCTCGTG
 GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTTGTCTTCTAGGATCAAC
 TCGGTCATTACCACAGCTCAAACCTGCCTTGGGACTCCCTCCCACAAAACCTGGCTCCGGATCAGG
 GAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTTCCTTCTTTAAGTCTGATACCATTAACA
 CAGATGCTCACACTGGGGCCAGATCTGCATCTGTAAATCCTGCTGCAGGAATGACACCTGGTAC
 CCAGACCCACCCATTGACCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
 CAATTTTTGTACACAACCTTGAGCCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC
 TTCACGAGCCTCATCATCCATTCTTGTTCCTGGGAGGCATCCTGCCACCAGTCAGGCAGGGGC
 TAATCCAGATGTCCAGGATGGAAGCCTTCCAGCAGGAGGAGCAGGTGTAAATCCTGCCACCCAGG
 GAACCCAGCAGGCCGCTCCCAACTCCCACTGGCAGATGACGACTTTGCAGTGACCACCCCT
 GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGGAATTCA
 GTAAAGCTGTTTCAAATTTTTTCACTAAGCTGCCTCGAATTTGGTGATACATGTGAATCTTTATC
 ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAATTCTTAATT
 TACCTGAAAATATTCTTGAAATTTAGAAAATATGTTCTATGTAGAGAATCCCAACTTTTAAAAA
 CAATAATTCAATGGATAAATCTGTCTTTGAAATATAACATTATGCTGCCTGGATGATATGCATAT
 TAAAACATATTTGAAAACTGAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM
LTLGPDHLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTTLSSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:
amino acids 1-16

FIGURE 97

SCTCAAGTGCCCTGCCTTGCCCCACCCAGCCCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCTCT
 CTCTCTTGCTTGGCAGCTGGACCAAGGGAGGCCAGTCTTGGGCGCTGGAGGGCCTGTCTTGACC**ATG**
 GTCCCTGCTTGGCTGTGGCTGCTTTGTGCTCTCCGTCCTCCCAAGGCCAGCCGTGC
 AGAGCTGTCTGTGAAGTTCCAGAAACATATGGTGAAATTTCCCTTTATACCTGACCAAGTTGC
 CGCTGCCCCGTGAGGGGGCTGAAGGCCAGATCGTGCTGTGAGGGGACTCAGGCAAGGCAACTGAG
 GGCCCATTTGCTATGGATCCAGATTCTGGCTTCTGCTGGTGACCAGGGCCCTGGACCGAGAGGA
 GCAGGCAGAGTACCAGCTACAGGTCACCCCTGGAGATCGAGGATGGACATGTCTTGTGGGGTCCAC
 AGCCTGTGCTTGTGACCGTGAAGGTAGAATTACCAAGTGCCCCATTTCTCTCAAGCCATCTA
 AGAGCTCGGCTGAGCCGGGGTACAGGCCTGGCATCCCCCTCCTCTTCTTTGAGGCTTCAGACCG
 GGATGAGCCAGGCACAGCCAACTCGGATCTTCGATTCCACATCTTGAGCCAGGCTCCAGCCCAG
 CTTCCCCAGACATGTTCCAGCTGGAGCCCTCGGCTGGGGGCTGTGGCCCTCAGCCCCAAGGGGAGC
 ACCAGCCTTGACCACGCCCTGGAGAGGACCTACCAGCTGTTGGTACAGGTCAGGGACATGGGTGA
 CAGGCCCTCAGGCCACCAAGGCCATGCCACCGTGAAGTCTCCATCATAGAGACCTGGGTGT
 CCCTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGTCTTATACCCGCACCACATGGCCCAGGTA
 CACTGGAGTGGGGGTGATGTGCATATCACCTGGAGAGCCATCCCCGGGACCCCTTTGAAGTGAA
 TGCAGAGGGAAACCTCTACGTGACCACAGAGCTGGACAGAGAAGCCACAGGCTGAGTACCTGCTCC
 AGGTGCGGGCTCAGAATTCCTATGGCAGGACTATCGGCCCTCTGGAGCTGCACGTGCTGGTG
 ATGGATGAGAATGACAACGTGCCATCTGCCCCTCCCCGTGACCCACAGTCAGCATGCTCCCTGAGCT
 CAGTCCACCAGGTACTGAAGTGACTAGACTGTGACGAGAGGATGCAGATGCCCCCGGCTCCCCCA
 ATTCCCACGTTGTGATCAGCTCTGAGCCCTGAGCCTGAGGATGGGTGAGGGGAGAGCCCTTC
 CAGGTGGACCCCACTTCAGGCAGTGTGACGCTGGGGTGCTCCCACCTCGAGCAGGCCAGAACAT
 CCTGCTTCTGGTGCCTGGCCATGGACCTGGCAGGCCAGAGGGTGGCTTCAGCAGCAGCTGTGAAG
 TCGAAGTCGCAGTCACAGATATCAATGATCACGCCCTGAGTTCATCACTTCCCAGATTGGGCCT
 ATAAGCCTCCCTGAGGATGTGGAGCCCGGAGCTCTGGTGCCATGCTAACAGCCATTGATGCTGA
 CCTCGAGCCGCCCTTCGCGCTCATGGATTGTGCCATTGAGAGGGGAGACACAGAAGGAGCACTTTG
 GCTTGATTGGGAGCCAGACTCTGGGCATTGTAGACTCAGACTTGAAGAACCTCAGTTATGAG
 GCAGCTCCAAGTCATGAGGTGGTGGTGGTGGTGCAGAGTGTGGCGAAGCTGGTGGGGCCAGGCC
 AGGCCCTGGAGCCACCGCCACGGTGACTGTGCTAGTGGAGAGAGTGATGCCACCCCCCAAGTTGG
 ACCAGGAGAGCTACCGAGGCCAGTGCCCCATCAGTGCCCCAGCCGGCTCTTTCTCTGCTGACCATC
 CAGCCCTCCGACCCCATCAGCCGACCCCTCAGGTTCTCCCTAGTCAATGACTCAGAGGCTGGCT
 CTGCATTGAGAAATTCTCCGGGAGGTTGCACACCGCCAGTCCCTGCAGGGCGCCAGCCTGGGG
 ACACCTACACGGTGCTTGTGGAGGCCAGGATACAGCCCTGACTCTTGCCCTGTGCCCTCCCAA
 TACCTCTGCACACCCCGCAAGACCATGGCTTGATCGTGAGTGGACCCGACGAAGGACCCCGATCT
 GGCCAGTGGGCACGGTCCCTACAGCTTCAACCCTTGGTCCCAACCCACGGTGCACGGGATTGGC
 GCCTCCAGACTCTCAATGGTTCCCATGCCTACCTACCTTGGCCCTGCATTGGGTGGAGCCAGT
 GAACACATAATCCCCGTGGTGGTCAGCCACAATGCCAGATGTGGCAGCTCCTGGTTCGAGTGAT
 CGTGCTCGCTGCAACGTGGAGGGGCAGTGCATGCGCAAGGTGGGCGCATGAAGGGCATGCCCA
 CGAAGCTGTGCGCAGTGGGCATCCTTGTAGGACCCCTGGTAGCAATAGGAATCTTCTCATCTCTC
 ATTTTCAACCCTGGACCATGTCAAGGAAGAAGACCCGGATCAACGACGACAGCGTGCCCT
 GAAGGCGACTGTCT**CGAA**TGGCCCAGGCAGCTCTAGCTGGGAGCTTGGCCTCTGGCTCCATCTGAG
 TCCCCTGGGAGAGAGCCAGCACCCAAGATCCAGCAGGGGACAGGACAGAGTAGAAGCCCTCCA
 TCTGCCCTGGGGTGGAGGACCATCACCATCACCAGGCATGTCTGCAGAGCCTGGACACCAACTT
 TATGGACTGCCATGGGAGTGCTCCAAATGTCAAGGTGTTGCCCAATAATAAGCCCCAGAGAA
 CTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 98

MVPAWLWLLCVSPQALPKAQPAELSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSGKAT
EGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVPHFSQAI
YRRLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQSPDMFQLEPRLGALALSPKG
STSLDHALERTYQLLVQVKDMGDQASGHQATATVEVSI IESTWVSLEPIHLAENLKVLYPHHMAQ
VHWSGGDVHYHLESHPPGPFENVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDIAAPLELHVL
VMDENDNVPICPPRDPTVSIPELSPPGTEVTRLAEDADAPGSPNSHVYQLLSPEPEDGVEGRA
FQVDPTSGSVTLGVLPLRAGQNILLVLAMDLAGAEGGFSSSTCEVEVAVTDINDHAPEFITSQIG
PISLPEDVEPGTLVAMLTADADLEPAFRLMDFAIERGDTEGTFGLDWEPDSGHVRLRLCKNLSY
EAAPSHEVVVVQSVAKLVGPGPGGATATVTVLVERVMPPPKLDQESYEASVPISAPAGSFLLT
IQPSDPISRTLRFSLVNDSEGWLCEIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALTALPVPS
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGPNPTVQRDWRLQTLNGSHAYLTALHWVEP
REHIIPVVVSHNAQMWQLLVRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGILVGTILVAIGIFLI
LIFTHWTMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCACGCCTG
 AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCCTTTA
 TCTCTTCACCTTCAAGTCCCTTTCTCAAGAATCCTCTGTTCTTTGCCCTCTAAAGTCTTGGTAC
 ATCTAGGACCCAGGCATCTTGCTTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG
 TTCTCCTTATGTTTGGTCTACTATTGCATTTAGAAGCTGCAACAAATTCATGAGACTAGCACC
 TCTGCCAACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCACCAACTCTGGGTCCAG
 TGTGACCTCCAGTGGGGTCAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCCAATGGGGTCA
 GCATAGTCACCAACTCTGAGTTCCATACAACTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
 TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGG
 GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCCTCCAGTGGGGCCAGCACAGTCCACCACT
 CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCC
 AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGGCCAGCACAGCCAC
 CAACTCTGACTCCAGCACAACTCCAGTGGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
 CCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACT
 GCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAG
 AACGACCTCCAATGGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGGCCA
 GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAG
 TCCAGCACGACCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG
 GGCTAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCGGCACAGCCACCAACT
 CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCCACCAATTCTGAGTCCAGCACACCTCC
 AGTGGGGCCAAACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGGCCAAACAGCCAC
 CAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAA
 CCTCCAGTGGGGTCAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCTAGCACAA
 GCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCTAG
 CACAGTGTCCAGTGGGATCAGCACAGTCCACCAATTCTGAGTCCAGCACAACTCCAGTGGGGCCA
 ACACAGCCACCAACTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGA
 ATGCACACAACTTCCCATAGTGCATCTACTGCAGTGAAGGCAAGCCTGGTGGGTCCCTGGT
 GCCGTGGGAAATCTTCTCATCACCTGGTCTCGGTTGTGGCGGCCGTGGGGCTCTTGCTGGGC
 TCTTCTTCTGTGTGAGAAACAGCCTGTCCCTGAGAAACACCTTTAACACAGCTGTCTACCACCT
 CATGGCCTCAACCATGGCCTTGGTCCAGGCCCTGGAGGGAATCATGGAGCCCCCAGAGGCCAG
 GTGGAGTCTAACTGGTTCTGGAGGAGACCAGTATCATCGATAGCCATGGAGATGAGCGGGAGGA
 ACAGCGGGCCCTGAGCAGCCCCGGAAGCAAGTGCCGCATTCTTCAGGAAGGAAGAGACCTGGGCA
 CCAAGACCTGGTTTCTTTCATTTCATCCAGGAGACCCCTCCAGCTTTGTTTGAGATCCTGAA
 AATCTTGAAGAAGGTATTCTCACCTTTCTTGCTTTACCAGACACTGGAAAGAGAATACTATAT
 TGCTCATTTAGCTAAGAAATAAATACATCTCATCTAACACACACGACAAAGAGAAGCTGTGCTTG
 CCCCGGGTGGGTATCTAGCTCTGAGATGAACCTAGTTATAGGAGAAAACCTCCATGCTGGACTC
 CATCTGGCATTCAAAATCTCCACAGTAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATISGS
SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESSTPSS
GASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTSSGASTA
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTNGAGTATNSES
STTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNSDSSTTSS
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
TNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSEASTATNSESSTVSSGISTVTNSES
STTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVSVVA
AVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGPGPGGNHGAPHRPRWSPNWFWRPVSII
AMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCTCCGCGTTACGGGATGAATTAACGGCGGGTTCCGCACGGAGGTTGTGACCCCTA
CGGAGCCCCAGCTTGCCACGCACCCCACTCGGCGTCGCGCGGCGTGCCCTGCTTGTCACAGGTG
GGAGGCTGGAACATATCAGGCTGAAAAACAGAGTGGGTACTCTCTTCTGGGAAGCTGGCAACAAAT
GGATGATGTGATATATGCATTCCAGGGGAAGGGAAATTGTGGTGCTTCTGAACCCATGGTCAATT
AACGAGGCAGTTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCTTTGGAATCAT
GGTGTGATGGAAAGGATTTACTTTTATACTGACTCTGTTTTGGGGAAGCTTTTTTGGGAAGCATTT
TCATGCTGAGTCCCTTTTTTACCTTTTGATGTTTGTAACCCATCTTGGTATCGCTGGATCAACAAC
CGCCTTGTTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCATGTTTGGTGTAAGT
GATTATAACTGGGGATGCATTTGTTCCCTGGAGAAAGAAGTGTCATTATCATGAACCATCGGACAA
GAATGGACTGGATGTTCCCTGTGGAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT
TGCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTTGGTTGGGCCATGCAGGCTGCTGCCTATAT
CTTCATTATAGGAAATGGAAGGATGACAAGAGCCATTTCGAAGACATGATTGATTACTTTTGTG
ATATTACGAACCACTTCAACTCCTCATATTCCCAGAAGGGACTGATCTCACAGAAAACAGCAAG
TCTCGAAGTAATGCATTTGCTGAAAAAATGGACTTCAGAAATATGAATATGTTTTACATCCAAG
AACTACAGGCTTTACTTTTGTGGTAGACCGTCTAAGAGAAGGTAAGAACCTTGATGCTGTCCATG
ATATCACTGTGGCGTATCCTCACAACATTCCTCAATCAGAGAAGCACCTCCTCCAAGGAGACTTT
CCCAGGGAAATCCACTTTCACGTCCACCGGTATCCAATAGACACCTCCCCACATCCAAGGAGGA
CCTTCAACTCTGGTGCCACAAACGGTGGGAAGAGAAAGAAGAGAGGCTGCGTTCCTTCTATCAAG
GGGAGAAGAATTTTTATTTTACCGGACAGAGTGTCATTCCACCTTGCAAGTCTGAACCTCAGGGTC
CTTGTTGGTCAAATTGCTCTCTATACTGTATTGGACCTGTTTACGCCCTGCAATGTGCCTACTCAT
ATATTTGTACAGTCTTGTTAAGTGGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGA
GAATATTTGGTGGACTGGAGATCATAGAACTTGCAATGTTACCGACTTTTACACAAACAGCCACAT
TTAAATTCAAAGAAAAATGAGTAAGATTATAAGGTTTGCCATGTGAAAACCTAGAGCATATTTTG
GAAATGTTCTAAACCTTTCTAAGCTCAGATGCATTTTTCATGACTATGTGCAATATTTCTTACT
GCCATCATTATTTGTTAAAGATATTTTGCACCTTAATTTTGTTGGGAAAAATATTGCTACAATTTT
TTTAATCTCTGAATGTAATTTTCGATACTGTGTACATAGCAGGGAGTGATCGGGGTGAAATAACTT
GGGCCAGAATATTATTAAACAATCATCAGGCTTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSF GIMVSWKGIYFILTLFWGSFFGSI FMLSP
FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM
FLWNCIMRYSYLRLKICLKASLKGVPFGFWAMQAAAYIFIHRKWDDKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSKSRSNFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDAVHDITVA
YPHNIPQSEKHLQGDFFPREIH FHVHRYPIDTLPTSKEDQLWCHKRWEEKERLRSFYQGEKNF
YFTGQSVIPPCKSELRLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG
LEIIE LACYRL LHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTGACC
 TCCAAATCATCCATCCACCCCTGCTGTCTATCTGTTTTCATAGTGTGAGATCAACCCACAGGAATA
 TCCATGGCTTTTGTGCTCATTTTGGTTCTCAGTTTCTACGAGCTGGTGTGAGGACAGTGGCAAGT
 CACTGGACCGGGCAAGTTTGTCCAGGCCTTGGTGGGGGAGGACGCCGTGTTCTCCTGCTCCCTCT
 TTCTGAGACCAGTGCAGAGGCTATGGAAGTGCGGTTCTCAGGAATCAGTTCCATGCTGTGGTC
 CACCTCTACAGAGATGGGGAAGACTGGGAATCTAAGCAGATGCCACAGTATCGAGGGAGAACTGA
 GTTTGTGAAGGACTCCATTGCAGGGGGGCGTGTCTCTCTAAGGCTAAAAACATCACTCCCTCGG
 ACATCGGCCTGTATGGGTGCTGGTTTCAGTTCCCAGATTACGATGAGGAGGCCACCTGGGAGCTG
 CGGGTGGCAGCACTGGGCTCACTTCCTCTCATTTCCATCGTGGGATATGTTGACGGAGGTATCCA
 GTTACTCTGCCTGTCCTCAGGCTGGTTCCCCAGCCACAGCCAAGTGGAAGGTCCACAAGGAC
 AGGATTTGTCTTCAGACTCCAGAGCAAAATGCAGATGGGTACAGCCTGTATGATGTGGAGATCTCC
 ATTATAGTCCAGGAAAATGCTGGGAGCATATTGTGTTCCATCCACCTTGCTGAGCAGAGTCATGA
 GGTGGAATCCAAGGTATTGATAGGAGAGACGTTTTTCCAGCCCTCACCTTGGCGCCTGGCTTCTA
 TTTTACTCGGGTACTCTGTGGTGCCCTGTGTGGTGTGTCATGGGGATGATAATTGTTTTCTTC
 AAATCCAAAGGGAAAATCCAGGCGGAACCTGGACTGGAGAAGAAAGCACGGACAGGCAGAATTGAG
 AGACGCCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCG
 TTTCTGATCTGAAAACCTGTAACCCATAGAAAAGCTCCCCAGGAGGTGCCCTCACTCTGAGAAGAGA
 TTTACAAGGAAGAGTGTGGTGGCTTCTCAGGGTTTCCAAGCAGGGAGACATTACTGGGAGGTGGA
 CGTGGGACAAAATGTAGGGTGGTATGTGGGAGTGTGTCGGGATGACGTAGACAGGGGGAAGAACA
 ATGTGACTTTGTCTCCCAACAATGGGTATTGGGTCTCTCAGACTGACAACAGAACATTTGTATTTT
 ACATTCAATCCCCATTTTATCAGCCTCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCCT
 GGACTATGAGGGTGGGACCATCTCCTTCTCAATACAAATGACCAGTCCCTTATTTATACCCTGC
 TGACATGTCTAGTTTGAAGGCTTGGTTGAGACCTTATATCCAGCATGCGATGTATGACGAGGAAAAG
 GGGACTCCCATATTTCATATGTCCAGTGTCTGGGGATGAGACAGAGAAGACCCTGCTTAAAGGGC
 CCCACACCACAGACCCAGACACAGCCAAGGGAGAGTGTCCCAGAGGTGGCCCCAGCTTCTCTCT
 CCGGAGCCTGCGCACAGAGAGTACGCCCCCCTCTCCTTTAGGGAGCTGAGGTTCTTCTGCCC
 TGAGCCCTGCAGCAGCGGCAGTCACAGCTTCCAGATGAGGGGGGATTGGCCTGACCCTGTGGGAG
 TCAGAAGCCATGGCTGCCCTGAAGTGGGGACGGAATAGACTCACATTAGGTTTAGTTTGTGAAAA
 CTCCATCCAGCTAAGCGATCTTGAACAAGTCACAACCTCCCAGGCTCCTCATTGTAGTCACGG
 ACAGTGATTCTTGCCTCACAGGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTTGCAGGTT
 TGAGGGCACAGTGTGTGCTAATGATGTGTTTTATATTATACATTTTCCACCATAAACTCTGTT
 TGCTTATTCCACATTAATTTACTTTTTCTATACCAAATCACCCATGGAATAGTTATTGAACACC
 TGCTTTGTGAGGCTCAAAGAATAAAGAGGAGGTAGGATTTTTCACTGATTCTATAAGCCCAGCAT
 TACCTGATACCAAACCAGGCAAAGAAAACAGAAGAAGAGGAAGGAAAACCTACAGGTCCATATCC
 CTCATTAAACACAGACACAAAATTTCTAAATAAAATTTTAAACAAATTAACATAAATATATTTA
 AAGATGATATATACTACTCAGTGTGGTTTGTCCCAAAATGCAGAGTTGGTTTAAATATTTAAAT
 ATCAACCAGTGTAAATTCAGCACATTAATAAAGTAAAAAAGAAAACCATAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVVH
LYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
VAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGDLSSDSRANADGYSLYDVEISI
IVQENAGSILCSIHLAEQSHEVESKVLIGETFFQSPWRLASILLGLLCGALCGVVMGMIIVFFK
SKGKIQAELDWRKKGQAEIRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF
TRKSVVASQGFQAGRHYWEVDVGQNVGWYVGVCRDDVDRGKNNVTLSPNNGYWVLRRLTEHLYFT
FNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLLRPYIQHAMYDEEKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTTCATTGCTGGTTGGCAATGATGTATCGGCCAGATGTGGTGAGGGCTAGGAAAAGAG
 TTTGTTGGGAACCCCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGTCTGGCAGTGTGCATTGGA
 CTCACTGTTTCATTATGTGAGATATAATCAAAAGAAGACCTACAATTACTATAGCACATTGTCATTTACAAC
 TGACAAACTATATGCTGAGTTTGGCAGAGAGGCTTCTAACAATTTTACAGAAATGAGCCAGAGACTTGAAT
 CAATGGTGAAAAATGCATTTTATAAATCTCCATTAAGGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTC
 AGTCAACAGAAGCATGGAGTGTGGCTCATATGCTGTTGATTGTAGATTCTACTCTACTGAGGATCCTGA
 AACTGTAGATAAAATTTGTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTAG
 ATCCTCACTCAGTTAAATTAATAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTGCTGCGGA
 ACACGAAGAAGTAAACTCTAGGTCAGAGTCTCAGGATCGTTGGTGGGACAGAAGTAGAAGAGGGTGAATG
 GCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCATCGCTGTGGAGCAACCTTAATTAATGCCACATGGC
 TTGTGAGTGTGCTCACTGTTTACAACATATAAGAACCCTGCCAGATGGACTGCTTCCTTTGGAGTAACA
 ATAAACCTTCGAAATGAAACGGGGTCTCCGGAGAATAATTGTCCATGAAAAATACAAACACCCATCACA
 TGACTATGATATTTCTCTTGACAGACTTTCTAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTTGTCT
 TCCCTGATGCATCCTATGAGTTTCAACCAGGTGATGTGATGTTTGTGACAGGATTTGGAGCACTGAAAAAT
 GATGGTTACAGTCAAAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAACCTGCAATGAACC
 TCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTTAGAAGGAAAAACAGATGCAT
 GCCAGGGTGACTCTGGAGGACCCTGGTTAGTTCAGATGCTAGAGATATCTGGTACCTTGCTGGAATAGTG
 AGCTGGGGAGATGAATGTGCGAAACCAACAAGCCTGGTGTATTAAGTACAGTTACGGCCTTGCGGGACTG
 GATTACTTCAAAACTGGTATCTAAGAGACAAAAGCCTCATGGAACAGATAACATTTTTTTTGTTTTTTG
 GGTGTGGAGGCCATTTTAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTGACTGATCTCA
 ATAAACTGTTTGCTTGATGCATGTATTTCTTCCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTGCCA
 GATCAACTCTGTCTATCTGTGAGCAATAGTTGAAACTTTATGTACATAGAGAAATAGATAATACAATATTAC
 ATTACAGCCTGTATTTCATTGTTCTCTAGAAGTTTGTGAGAATTTGACTTGTGACATAAATTTGTAAT
 GCATATATACAATTTGAAGCACTCCTTTCTTCAGTTTCTCAGCTCCTCTCATTTCAGCAAATATCCATTT
 TCAAGGTGCAGAACAGGAGTGAAAGAAAATATAAGAAGAAAAAATCCCTACATTTTATTTGGCACAGAA
 AAGTATTAGGTGTTTTCTTAGTGAATATTAGAAATGATCATATTCATTATGAAAGGTCAAGCAAAGACA
 GCAGAATACCAATCACTTCATCATTTAGGAAGTATGGGAACCTAAGTTAAGGAAGTCCAGAAAGAAGCCAAG
 ATATATCCTTATTTTCAATTCACAACTACTATGATAAATGTGAAGAAGATTCTGTTTTTTGTGACCT
 ATAATAATTATACAACTTCATGCAATGTACTTGTCTAAGCAAATTAAGCAAATATTTATTTAACATTG
 TTACTGAGGATGTCAACATATAACAATAAATATAAATCACCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTYNYSTLSFTTDKLY
 AEFGREASNNFTEMSQRLESMVKNAFYKSPLREEFVKSQVIKFSQQKHGVLAHMLLICRFHSTED
 PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG
 TEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIKPSKMKRGL
 RRIIVHEKYKHPSHDYDISLAELSSPVPTYNAVHRVCLPDASYEFQPGDVMFVTGFGALKNDGYS
 QNHLRQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGDGGPLVSSDARDIWYLAG
 IVSWGDECAKPNKPGVYTRVTALRDWITSKTI

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTTCCTG
 CCCCAGTGAAGCCCCCGCGTCCCGACTATCCCCAGGCGGGCGTGGGGCACCGGGCCCAGC
 GCCGACGATCGCTGCCGTTTTGCCCTTGGGAGTAGGATGTGGTGAAAGGATGGGGCTTCTCCCTT
 ACGGGGCTCACAATGGCCAGAGAAGATTCCGTGAAGTGTCTGCGCTGCCTGCTCTACGCCCTCAA
 TCTGCTCTTTTGGTTAATGTCCATCAGTGTGTTGGCAGTTTCTGCTTGGATGAGGGACTACCTAA
 ATAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTTTGACTTACTTTCTCT
 GTGGTTCATCCGGTCATGATTGCTGTTTGTCTGTTTCTTATCATTGTGGGGATGTTAGGATATTG
 TGGAACGGTGAAAAGAAATCTGTTGCTTCTTGCATGGTACTTTGGAAGTTTGCTTGTCAATTTTCT
 GTGTAGAACTGGCTTGTGGCGTTTGGACATATGAACAGGAACCTTATGGTTCAGTACAATGGTCA
 GATATGGTCACTTTGAAAGCCAGGATGACAAATTATGGATTACCTAGATATCGGTGGCTTACTCA
 TGCTTGGAAATTTTTTTCAGAGAGAGTTAAGTGCTGTGGAGTAGTATATTTCACTGACTGGTTGG
 AAATGACAGAGATGGACTGGCCCCAGATTCTGCTGTGTAGAGAATTTCCAGGATGTTCCAAA
 CAGGCCCACCAGGAAGATCTCAGTGACCTTTATCAAGAGGGTTGTGGGAAGAAATGTATTCCTT
 TTTGAGAGGAACCAACAACCTGCAGGTGCTGAGGTTTCTGGGAATCTCCATTGGGGTGACACAAA
 TCCTGGCCATGATTCTCACCATTACTCTGCTCTGGGCTCTGTATTATGATAGAAGGGAGCCTGGG
 ACAGACCAAATGATGTCTTGAAGAATGACAACTCTCAGCACCTGTCTATGTCCCTCAGTAGAACT
 GTTGAAACCAAGCCTGTCAAGAATCTTTGAACACACATCCATGGCAAACAGCTTTAATACACACT
 TTGAGATGGAGGAGTTATATAAAGAAATGTCACAGAAGAAAACCACAACTTGTTTTATTGGACT
 TGTGAATTTTGGAGTACATCTATGTGTTTCAGAAATATGTAGAAATAAAAATGTTGCCATAAAA
 TAACACCTAAGCATATACTATTCTATGCTTTAAATGAGGATGGAAAAGTTTCATGTCATAAGTC
 ACCACCTGGACAATAATTGATGCCCTTAAATGCTGAAGACAGATGTCATACCCACTGTGTAGCC
 TGTGTATGACTTTTACTGAACACAGTTATGTTTTGAGGCAGCATGGTTTGATTAGCATTTCCGCA
 TCCATGCAACGAGTCACATATGGTGGGACTGGAGCCATAGTAAAGTTGATTTACTTCTACCAA
 CTAGTATATAAAGTACTAATTAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTATTA
 CTCAGCGATCTATTCTTCTGATGCTAAATAAATTATATATCAGAAAACCTTTCAATATTGGTGACT
 ACCTAAATGTGATTTTTGCTGGTTACTAAAATATTCTTACCACCTAAAAGAGCAAGCTAACACAT
 TGTCTTAAGCTGATCAGGGATTTTTGTATATAAGTCTGTGTTAAATCTGTATAATTCAGTCGAT
 TTCAGTTCTGATAATGTTAAGAATAACCATTATGAAAAGGAAAATTTGTCCTGTATAGCATCATT
 ATTTTTAGCCTTTCTGTTAATAAAGCTTTACTATTCTGTCTGGGCTTATATTACACATATAAC
 TGTATTTTAAATACTTAACCACTAATTTTGAAAATTACCAGTGTGATACATAGGAATCATTATTC
 AGAATGTAGTCTGGTCTTTAGGAAGTATTAATAAGAAAATTTGCACATAACTTAGTTGATTGAGA
 AAGGACTTGTATGCTGTTTTTCTCCAAATGAAGACTCTTTTGACACTAAACACTTTTAAAAA
 GCTTATCTTTGCCTTCTCCAAACAAGAAGCAATAGTCTCCAAGTCAATATAAATTTCTACAGAAA
 TAGTGTTCTTTTCTCCAGAAAAATGCTTGTGAGAATCATTAAACATGTGACAATTTAGAGATT
 CTTTGTTTTATTTCACTGATTAATATACTGTGGCAAATTACACAGATTATTAATTTTTTTACAA
 GAGTATAGTATATTTATTTGAAATGGGAAAAGTGCATTTTACTGTATTTTGTGTATTTTGTAT
 TTCTCAGAATATGGAAGAAAATTAATGTGTCAATAAATATTTTCTAGAGAGTAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTTLTAETRVEEAVILTYFPVVHP
VMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMT
LKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTMDWPPDSCCVREFPGCSKQAHQ
EDLSDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWALYYDRREPQTDQM
MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTATCCCACCTCATCCTCTTCCTCTGATAAAGCCCCCTACCAGTGCT
 GATAAAGTCTTTCTCGTGAGAGCCTAGAGGCCTTAAAAAAGTGCTTGAAAGAGAAGGGGACAAAGGAACA
 CCAGTATTAAGAGGATTTCCAGTGTTTCTGGCAGTTGGTCCAGAAGGATGCCCTCCATTCTGCTTCTCACCTG
 CCTCTTCATCACAGGCACCTCCGTGTACCCGTGGCCCTAGATCCTTGTTCTGCTTACATCAGCCTGAATGAGC
 CCTGGAGGAACACTGACCACCAGTTGGATGAGTCTCAAGGTCCTCTCTATGTGACAACCATGTGAATGGGAG
 TGGTACCACCTTCACGGGCATGGCGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCACTGTGGAACCCA
 CGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAAGGCACGGCATTGTGCAACGCCAGGCTTGTGCCAGCT
 TCAATGGGAACGTCTGTCTCTGGAACACCACGGTGGAGTCAAGGCTTGCCCTGGAGGCTACTATGTGTATCGT
 CTGACCAAGCCCAGCGTCTGCTTCCACGTCTACTGTGGTCATTTTTATGACATCTGCGACGAGGACTGCCATGG
 CAGCTGCTCAGATACCAGCGAGTGCACATGCGCTCCAGGAAGTGTGCTAGGCCCTGACAGGCAGACATGCTTTG
 ATGAAAATGAATGTGAGCAAAACAACGGTGGCTGCACTGAGATCTGTGTGAACCTCAAAAACCTCTACCGCTGT
 GAGTGTGGGGTTGGCCGTGTGCTAAGAGTGATGGCAAGACTTGTGAAGACGTTGAAGGATGCCACAATAACAA
 TGGTGGCTGCAGCCACTCTTGCCTTGGATCTGAGAAAGGCTACCAAGTGTGAATGTCCCCGGGGCTGGTGTGT
 CTGAGGATAACCACACTTGCCAAGTCCCTGTGTGTGCAAAATCAAATGCCATTGAAGTGAACATCCCCAGGGAG
 CTGGTTGGTGGCCTGGAGCTCTTCTGACCAACACCTCCTGCCGAGGAGTGTCCAACGGCACCCATGTCAACAT
 CCTCTTCTCTCTCAAGACATGTGGTACAGTGGTCGATGTGGTGAATGACAAGATTGTGGCCAGCAACCTCGTGA
 CAGGTCTACCCAAGCAGACCCCGGGGAGCAGCGGGGACTTCATCATCCGAACAGCAAGCTGCTGATCCCGGTG
 ACCTGCGAGTTTCCACGCCTGTACACCATTTCTGAAGGATACGTTCCCAACCTTCGAAACTCCCCACTGGAAAT
 CATGAGCCGAAATCATGGGATCTTCCCATTCACTCTGGAGATCTTCAAGGACAATGAGTTTGAAGAGCCTTACC
 GGGAAAGCTCTGCCCACCCTCAAGCTTCTGACTCCCTCTACTTTGGCATTGAGCCCGTGGTGCACGTGAGCGGC
 TTGAAAAGCTTGGTGGAGAGCTGCTTTGCCACCCCCACCTCCAAGATCGACGAGGTCTGAAATACTACCTCAT
 CCGGGATGGCTGTGTTTCAGATGACTCGGTAAAGCAGTACACATCCCGGGATCACCTAGCAAAGCACTTCCAGG
 TCCCTGTCTTCAAGTTTGTGGGCAAGACCACAAGGAAGTGTTCCTGCACTGCCGGGTCTTGTCTGTGGAGTG
 TTGGACGAGCGTTCCCGCTGTGCCCAGGTTGCCACCGGCGAATGCGTCGTGGGGCAGGAGGAGGACTCAGC
 CGGTCTACAGGGCCAGACGCTAACAGGCGGCCCGATCCGCATCGACTGGGAGGACTAGTTTCGTAGCCATACCTC
 GAGTCCCTGCATTGGACGGCTCTGCTCTTTGGAGCTTCTCCCCCACCGCCCTCTAAGAACATCTGCCAACAGC
 TGGGTTTCAGACTTCACACTGTGAGTTTCAGACTCCACGACCAACTCACTCTGATTCTGGTCCATTCACTGGGCA
 CAGGTACAGCACTGCTGAACAATGTGGCCTGGGTGGGGTTTCATCTTTCTAGGGTTGAAAACCTAACTGTCCA
 CCCAGAAAGACACTCACCCATTTCCTCATTTCTTTCTTACACTTAAATACCTCGTGTATGGTGAATCAGAC
 CACAAAATCAGAAGCTGGGTATAATATTCAAGTTACAAACCTAGAAAAATTAACAGTTACTGAAATTATGA
 CTTAAATACCAATGACTCCTTAAATATGTAAATTATAGTTTACCTTGAATTTCAATTCAAATGCAGACTAA
 TTATAGGGAATTTGGAAGTGTATCAATAAACAGTATATAATTTT

2

Important features of the protein:

amino acids 1-16

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

amino acids 411-418, 443-451

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

amino acids 197-209

amino acids 431-457

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCCTGCACTCGG
 GCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACCTGTGTGGGGAGGCCCT
 CCTGCTGCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCATGT
 TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAAACCCCGTATCCCC
 ATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATATAGCACTACTGAGCCTGGCGAGTATCATCATTGT
 GGTGTCTCATCAAGGTGATTCTGGATAAATACTACTTCCTCTGCGGGCAGCCTCTCCACTTCATCCCGA
 GGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC
 CCCGAAGGGCCTGCAGTGGCAGTCCGCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC
 AGGGAACCTGGTTCTCTGCCTGTTTCGACAACCTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGATGG
 GCTACAGCAGAGCTGTGGAGATTGGCCCAGACCAGGATCTGGATGTTGTTGAAATCACAGAAAACAGCCAG
 GAGCTTCGATGCGGAACCAAGTGGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCACGTGCTTTCCTG
 TGGGAAGAGCCTGAAGACCCCCGTGTGGTGGGTGGGGAGGAGCCTCTGTGGATTCTTGGCCTTGGCAGG
 TCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCTCACGGCA
 GCCCCTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAACTGGGCAG
 CTTCCCATCCCTGGCTGTGGCCAAGATCATCATTTGAATTCAACCCCATGTACCCCAAAGACAATGACA
 TCGCCCTCATGAAGCTGCAGTTCCCACTCACTTTCTCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTT
 GATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTTGGATGGGGCTTACGAAGCAGAATGGAGGAA
 GATGTCTGACATACTGTGTCAGGCGTCAGTCCAGGTCATTGACAGCACACGGTGCAATGCAGACGATGCGT
 ACCAGGGGGAAGTCACCGAGAAGATGATGTGTGTCAGGCATCCCGGAAGGGGGTGTGGACACCTGCCAGGGT
 GACAGTGGTGGGCCCCTGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGG
 CTGCGGGGGCCCCGAGCACCCAGGAGTATACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCT
 GGAAGGCTGAGCTGTAATGCTGCTGCCCTTTGCAGTGTCTGGGAGCCGCTTCCTTCCTGCCCTGCCACCT
 GGGGATCCCCCAAAGTCAGACACAGAGCAAGAGTCCCCCTGGGTACACCCCTCTGCCACAGCCTCAGCAT
 TTCTTGAGCAGCAAAGGGCCTCAATTCTGTAAAGAGACCTCGCAGCCCAGAGGCGCCCAGAGGAAGTCA
 GCAGCCCTAGCTCGGCCACACTTGGTGCTCCCAGCATCCAGGGAGAGACACAGCCCACTGAACAAGGTCT
 CAGGGGTATTGCTAAGCCAAGAAGGAACCTTCCCACACTACTGAATGGAAGCAGGCTGTCTTGTAAGGCC
 CAGATCACTGTGGGCTGGAGAGGAGAAGGAAAGGCTCTGCGCCAGCCCTGTCCGTCTTACCCATCCCCAA
 GCCTACTAGAGCAAGAAACAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTTACCT
 ACTGTTGTCTGTTATTACAGCTATGGCCACTATTATTAAAGAGCTGTGTAACATCTCTGGCAAAAAAA
 AAAA

FIGURE 112

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPIIIALSLASIIIVVLIKVILDKYYFLCG
 QPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDN
 FTEALAETACRQMGYSRAVEIGPDQDLDVVEITENSQELMRNSSGPCLSGSLVSLHCLACGKSL
 KTPRVVGGEASVDSWPWQVSIQYDKQHVCSSILDPHWVLTAAHCFRKHTDVFNWKVRAGSDKL
 GSFPSSLAVAKIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEELTPATPLWIIGWG
 FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTCQGDSGGPLMYQS
 DQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACCTCCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGATTA
TAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACTTTTTCAGCAACTAAAAAGCCACAGGAGT
TGAAGTGTCTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTACATTAAAATC
TGTTTTTTGTTCTCTTGTAAGTACCTTTACCTTCCTAACACAGAGGATCTGTCACTGTGGCTCT
GGCCCAAACCTGACCTTCACCTCTGGAACGAGAACAGAGGTTTCTACCCACACCGTCCCCTCGAAG
CCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTCACCAACTGTCTCACGTCT
GGAGGCACTGACTCGGGCAGTGACAGGTAGCTGAGCCTCTGGTAGCTGCGGCTTTCAAGGTGGGC
CTTGCCCTGGCCGTAGAAGGGATTGACAAGCCCGAAGATTTCATAGGCGATGGCTCCCACTGCCC
AGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGGCCAGGACGGGCCGTGGACACCTGCTCA
GAAGCAGTGGGTGAGACATCACGCTGCCCCGCCATCTAACCTTTTCATGTCTGCACATCACCTG
ATCCATGGGCTAATCTGAACTCTGTCCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC
CAGAAGGGGTCTGCTTAGACCACCTGGTTTATGTGACAGGACTTGCACTCTCTGGAACATGAGG
GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAGGAACCTGTGCCAAATTATGGGTGAGAAAAGATG
GAGGTGTTGGGTATCACAAAGGCATCGAGTCTCCTGCATTCACTGGACATGTGGGGGAAGGGCTG
CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGGCCATCAGACAGCCGTTTCCGCCCCGAT
CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAAA
TCTGCGATCACCAGCCAGGGGCAGCCGTCTGGGAAGGAGCAAGCAAAGTGACCATTCTCTCTCC
CTCCTTCCCTCTGAGAGGCCCTCCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGG
GGCTAATGGCTCAGTGTGGGCCAGGAGGTGAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCT
GTGCGAACACGAAATGCCTCCAGTAAGCACAGGCTGCAAAATCCCCAGGCAAAGGACTGTGTGG
CTCAATTTAAATCATGTTCTAGTAATTGGAGCTGTCCCAAGACCAAAGGAGCTAGAGCTTGTT
CAAATGATCTCCAAGGGCCCTTATACCCAGGAGACTTTGATTTGAATTTGAAACCCCAAATCCA
AACCTAAGAACCAGGTGCATTAAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATGCCAACAT
TTTGGGAGGCCGAGGCGGGTAGATCACCTGAGGTGAGGAGTTCAAGACCAGCCTGGCCAACATGG
TGAAACCCCTGTCTCTACTAAAAATACAAAAAACTAGCCAGGCATGGTGGTGTGTGCCTGTATC
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGTGAAGGAGGCTGAGACA
GGAGAATCACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAATAAAAAAGAATTA
TGTTTATTTGTAA

FIGURE 115

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA
AGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCCAAGAAA
ATATGTAAATCACTTAAGATTTGTGGACTGGTGTGTTGGTATCCTGGCCCTAACTCTAATTGTCCT
GTTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAAGCCTATGACATGGAGCACACTT
TCTACAGCAATGGAGAGAAGAAGAAGATTACATGGAAATTGATCCTGTGACCAGAACTGAAATA
TTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAACGGATACACTGG
CATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAGTGATTCTGAATTTT
CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCACAACTTCTTTGAACAGTCAGTG
ATTTGGGTCCCAGCAGAAAAAGCCTATTGAAAACCGAGATTTTCTTAAAAATTCCAAAATCTGGA
GATTTGTGATAACGTGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAG
ACTTTGAGGAGGAGGAGAGAAGATCTTCACTTTCTGCCAACGAAAAAAGGGATTGAACAAAAT
GAACAGTGGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGA
AGAACTTCCAATAAATGACTATACTGAAAATGGAATAGAATTTGATCCCATGCTGGATGAGAGAG
GTTATTGTTGATTTTACTGCCGTCGAGGCAACCGCTATTGCCGCCGCTCTGTGAACCTTTACTA
GGCTACTACCCATATCCATACTGCTACCAAGGAGGACGAGTCATCTGTCGTGTCATCATGCCTTG
TAACTGGTGGGTGGCCCGCATGCTGGGGAGGGTCTTAATAGGAGGTTTGAGCTCAAATGCTTAAAC
TGCTGGCAACATATAATAAATGCATGCTATTCAATGAATTTCTGCCATGAGGCATCTGGCCCT
GGTAGCCAGCTCTCCAGAATTACTTGTAGGTAATTCCTCTCTTCATGTTCTAATAAACTTCTACA
TTATCACCAAAAAAAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFKSKKICKSLKICGLVFGILALTLIVLFWGSKHFWPEVPKKAYDME
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLVHDFKNGYTGIIYFVGLQKCFIKTQIKVIP
EFSEPEEEIDENEEITTTFFEQSVIWPVPAEKPIENRDFLKNKILEICDNVTMYWINPTLISVSE
LQDFEEEGEDLHFPANEKKGIEQNEQWVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMLD
ERGYCCIIYCRRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-
242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCCCTCAGGAGCGGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCGCAGGCGGCA
 GGGCGGGCGGCCAGGATCATGTCCACCACCACATGCCAAGTGGTGGCGTTCTCCTGTCCATCCTGGGGCT
 GGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCAGGACCTGTACGACAACCCCGTCACCT
 CCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTCAGGCTTCACCGAATGCAGGCC
 TATTTACCATCCTGGGACTTCCAGCCATGCTGCAGGCAGTGCAGCCCTGATGATCGTAGGCATCGTCCT
 GGGTGCCATTGGCCTCCTGGTATCCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTG
 CCAAAGCCAACATGACACTGACCTCCGGGATCATGTTTCATTGTCTCAGGTCTTTGTGCAATTGCTGGAGTG
 TCTGTGTTTGCCAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGG
 GATGGTGACAGTGTTCAGACCAGGTACACATTTGGTGGCGCTCTGTTCTGGGGCTGGGTCTGGAGGCC
 TCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGGCCTGGCACCAGAAGAAACCAACTACAAA
 GCCGTTTCTTATCATGCCTCAGGCCACAGTGTTCCTTACAAGCCTGGAGGCTTCAAGGCCAGCACTGGCTT
 TGGGTCCAACACCAAAAACAAGAAGATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATC
 CTTCCAAGCAGCACTATGTGTAATGTCTTAAGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCA
 CCAAAAAACAAGGAGATCCCCTAGATTTCTTCTTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGCCT
 CGATTTTCATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCACCATAAAACA
 GCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTTCTTTTTTAAATATAACT
 TTCTACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCACATTTTGATGATTTAGACAGACTCCCCCTC
 TTCCTCTAGTCAATAAACCCATTGATGATCTATTTCCAGCTTATCCCCAAGAAAACCTTTGAAAGGAAA
 GAGTAGACCCAAAGATGTTATTTTCTGCTGTTTGAATTTTGTCTCCCCACCCCAACTTGCGTAGTAATAA
 ACACTTACTGAAGAAGAAGCAATAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTT
 ACACTGTGATCTTAAAAGTTACCAAACCAAAGTCATTTTCAGTTTGAGGCAACCAACCTTTCTACTGCTG
 TTGACATCTTCTTATTACAGCAACACCATTCTAGGAGTTTCCTGAGCTCTCCACTGGAGTCCTCTTTCTGT
 CGCGGGTCAGAAATTGTCCCTAGATGAATGAGAAAATTATTTTTTTTAAATTAAGTCCTAAATATAGTTAA
 AATAAATAATGTTTTAGTAAAATGATACACTATCTCTGTGAAATAGCCTCACCCCTACATGTGGATAGAAG
 GAAATGAAAAAATAATTGCTTTGACATTGTCTATATGGTACTTTGTAAAGTCATGCTTAAGTACAAATTC
 ATGAAAAGCTCACACCTGTAATCCTAGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAGAAGT
 TCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAAATCAGCCAGTCA
 TGGTGGCATACACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTGGGAGGATCACTTGAGCCCAGGGAGGT
 TGGGGCTGCAGTGAGCCATGATCACCACTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAA
 AATAAAAAATAAATAATGGAACACAGCAAGTCCTAGGAAGTAGGTAAACTAATTCCTTTAA

FIGURE 118

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLWRSCVRQSSGFTECRP
 YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGL
 CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA
 CRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQSYPSKHDY
 V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAACTGTTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCGGGAGTCC
 AGCTGGCTAAAACTCATCCCAGAGGATAAATGGCAACCCATGCCTTAGAAATCGCTGGGCTGTTTCTTG
 GTGGTGTGGAATGGTGGGCACAGTGGCTGTCACTGTCATGCCTCAGTGGAGAGTGTGGCCTTCATT
 GAAAACAACATCGTGGTTTTTGAAAACTTCTGGGAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA
 CATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTTCTCCGGACCTACAGGCAGCCAGAG
 GACTGATGTGTGCTGCTTCCGTGATGTCCTTCTTGCTTTTCATGATGGCCATCCTTGGCATGAAATGC
 ACCAGGTGCACGGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTTCAT
 CATCACGGGCATGGTGGTGTCTATCCCTGTGAGCTGGGTTGCCAATGCCATCATCAGAGATTTCTATA
 ACTCAATAGTGAATGTTGCCCAAAAACGTGAGCTTGGAGAAGCTCTCTACTTAGGATGGACCACGGCA
 CTGGTGTGATTGTTGGAGGAGCTCTGTTCTGCTGCGTTTTTTGTTGCAACGAAAAGAGCAGTAGCTA
 CAGATACTCGATACCTTCCCATCGCACAAACCAAAAAAGTTATCACACCGGAAAGAAGTCACCGAGCG
 TCTACTCCAGAAGTCAGTATGTGTAGTTGTGTATGTTTTTTAACTTTACTATAAAGCCATGCAAATG
 ACAAAAATCTATATTACTTTCTCAAAATGGACCCAAAGAAACTTTGATTACTGTTCTTAACTGCCT
 AATCTTAATTACAGGAACTGTGCATCAGCTATTTATGATTCTATAAGCTATTTACAGCAGAATGAGATA
 TTAAACCAATGCTTTGATTGTTCTAGAAAGTATAGTAATTTGTTTTCTAAGGTGGTTCAAGCATCTA
 CTCTTTTTATCATTACTTCAAAATGACATTGCTAAAGACTGCATTATTTTACTACTGTAATTTCTCC
 ACGACATAGCATTATGTACATAGATGAGTGTAACATTTATATCTCACATAGAGACATGCTTATATGGT
 TTTATTTAAAAATGAAATGCCAGTCCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGGAA
 ATCATGGATAGGGTTGAAGAAGGTACTATTAATTGTTTAAAAACAGCTTAGGGATTAATGTCCTCCA
 TTTATAATGAAGATTAAAAATGAAGGCTTAAATCAGCATTGTAAGGAAATTGAATGGCTTTCTGATAT
 GCTGTTTTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTTATCCTCTTCTCCCAGAGGCTTTTTTTTT
 CTTGTGTATTAAATTAACATTTTAAAAACGCAGATATTTGTCAAGGGGCTTGCATTCAAACCTGCTT
 TTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGATGGTTTTAGGAAAGTG
 AAAATATTTTTGTTTTGTATTGTAAGAAGAATGATGCATTTTGACAAGAAATCATATATGTATGGAT
 ATATTTTAATAAGTATTTGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAGAGCAGAAAAATA
 TGTCTTGGTTTTTCATTGCTTACCAAAAAACAACAACAAAAAAGTTGCCTTTGAGAACTTCACTT
 GCTCCTATGTGGGTACCTGAGTCAAAATGTGATTTTTGTTCTGTGAAAAATAAATTTCTTCTGTGA
 CCATTTCTGTTTAGTTTACTAAAAATCTGTAAATACTGTATTTTTCTGTTATTCCAAATTTGATGAA
 ACTGACAATCCAATTTGAAAGTTTGTGTCGACGTCTGTCTAGCTTAAATGAATGTGTTCTATTGCTT
 TATACATTTATATTAATAAATGTACATTTTCTAATT

FIGURE 120

MATHALEIAGLFLGGVMVGTVAVTVMPQWRVSAFIENNIVVFENFWEGLMNCVRQANIRMQCK
IYDSLALSPDLQAARGLMCAASVMSFLAFMAILGMKTRCTGDNEKVKAHILLTAGIIFIITG
MVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCCNEKSSSY
RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCGCGGGTGAAAGGCGCATTGATGCAGCCTGCGGGCGCCTCGGAGCGCGGCGGAG
CCAGACGCTGACCACGTTCTCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCGCTGCCCGGCAGCC
GGGAGCCCATGCAGACCCAGGGCCCCGCCGCTCCCCGAGCGGCTCCGCGGCCCTCTGCTGCTCC
TGCTGCTGCAGCTGCCCCGCGCGTCGAGCGCCTCTGAGATCCCCAAGGGGAAGCAAAAGGCGCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC
TGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCCGGGTACACCTGGGATCCCAGGTCTGGGATG
GATTCAAAGGAGAAAAGGGGAATGTCTGAGGGAAGCTTTGAGGAGTCCTGGACACCCAACCTAC
AAGCAGTGTTTCATGGAGTTCATTGAATTATGGCATAGATCTTGGGAAAATTGCGGAGTGACATT
TACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGTTTCAGTGGCTCACTTCGGCTAAAATGCA
GAAATGCATGCTGTCAGCGTTGGTATTTACATTCAATGGAGCTGAATGTTTCAGGACCTCTTCCC
ATTGAAGCTATAATTTATTGGACCAAGGAAGCCCTGAAATGAATTCAACAATTAATATTCATCG
CACTTCTTCTGTGGAAGGACTTTGTGAAGGAATTGGTGCTGGATTAGTGGATGTTGCTATCTGGG
TTGGCATTGTTTCAGATTACCCAAAAGGAGATGCTTCTACTGGATGGAATTCAGTTTCTCGCATC
ATTATTGAAGAACTACCAAAATAAATGCTTTAATTTTCATTGCTACCTCTTTTTTTATTATGCC
TTGGAATGGTTCACTTAAATGACATTTTAAATAAGTTTATGTATACATCTGAATGAAAAGCAAAG
CTAAATATGTTTACAGACCAAAGTGTGATTTACACTGTTTTTAAATCTAGCATTATTCTTTTG
CTTCAATCAAAGTGGTTTCAATATTTTTTTTAGTTGGTTAGAATACTTTCTTCATAGTCACATT
CTCTCAACCTATAATTTGGAATATTGTTGTGGTCTTTTGTTTTTTCTCTTAGTATAGCATTTTTTA
AAAAAATATAAAAGCTACCAATCTTTGTACAATTTGTAAATGTTAAGAATTTTTTTTATATCTGT
TAAATAAAAATTATTTCCAACA

FIGURE 122

MRPQGPAASPQRLRGLLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAAGVPGR
DGSPGANVIPGTPGIPGRDGFKEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK
MRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECSGPLPIEAIYLDQGSPENSTINIHRTS
SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCTGAA
 CTGGGTGCTCATCACGGGAACCTGCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
 ATTGCCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTTTTTTTA
 ACCGCCCCCTCCCCACCCCCAAAAAACTGTAAAGATGCAAAAACGTAATATCCATGAAGATCC
 TATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATTTATTTGTTCTTGAG
 TGTTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCTCCAAGGGGTCCAATTTT
 TCTTCTGGGTGTGAGGAGCCCTGACTCACTACAGTGCAGCTGACAGGGGCTGTCATGCAACTG
 GCCCCTAAGCCAAAGCAAAAGACCTAAGGACGACCTTTGAACAATACAAAGGATGGGTTTCAATG
 TAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCCACTGTCTTACTGACAATG
 CTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTGTAGGTGTGAAGGCAAAATGGTATATTGTGA
 ATCTCAGAAATTACAGGAGATACCCCAAGTATATCTGCTGGTTGCTTAGGTTTGTCCCTTCGCT
 ATAACAGCCTTCAAAAACCTAAGTATAATCAATTTAAAGGGCTCAACCAGCTCAGCTGGCTATAC
 CTTGACCATAACCATATCAGCAATATTGACGAAAATGCTTTTAAATGGAATACGCAGACTCAAAGA
 GCTGATTCTTAGTTCCAATAGAATCTCCTATTTTCTTAACAATACCTTCAGACCTGTGACAAAT
 TACGGAACCTGGATCTGTCTATAATCAGCTGCATTCTCTGGGATCTGAACAGTTTCGGGGCTTG
 CGGAAGCTGCTGAGTTTACATTTACGGTCTAACTCCCTGAGAACCATCCCTGTGCGAATATTCCA
 AGACTGCCGCAACCTGGAACTTTTGGACCTGGGATATAACCGGATCCGAAGTTTAGCCAGGAATG
 TCTTTGCTGGCATGATCAGACTCAAAGAACTTCACCTGGAGCACAATCAATTTTCCAAGCTCAAC
 CTGGCCCTTTTCCAAGGTTGGTCAGCCTTCAGAACCCTTACTTGCAAGTGAATAAAATCAGTGT
 CATAGGACAGACCATGTCTGGACCTGGAGCTCCTTACAAAGGCTTGATTTATCAGGCAATGAGA
 TCGAAGCTTTCAGTGGACCCAGTGTTTTCCAGTGTGTCCCGAATCTGCAGCGCTCAACCTGGAT
 TCCAACAAGCTCACATTTATGGTCAAGAGATTTTGGATTCTTGGATATCCCTCAATGACATCAG
 TCTTGCTGGGAATATATGGGAATGCAGCAGAAATATTGCTCCCTTGTAACCTGGCTGAAAAGTT
 TTAAAGGTCTAAGGGAGAATAACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAAATGTG
 ATCGATGCAGTGAAGAACTACAGCATCTGTGGCAAAAGTACTACAGAGAGGTTTGATCTGGCCAG
 GGCTCTCCCAAAGCCGACGTTTAAAGCCCAAGCTCCCGAGGCCGAAGCATGAGAGCAAAACCCCTT
 TGCCCCGACGGTGGGAGCCACAGAGCCCGGCCAGAGACCGATGCTGACGCCGAGCACATCTCT
 TTCCATAAAATCATCGCGGGCAGCGTGCGCTTTTCTGTCCGTGCTCGTCATCTGCTGTTAT
 CTACGTGTCTATGGAAGCGGTACCCTGCGAGCATGAAGCAGCTGCAGCAGCGCTCCCTCATGCGAA
 GGCACAGGAAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCCAGCACCAGGAATTTTATGTA
 GATTATAAACCCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGGACGGGACCCTGCACCTA
 TAACAAATCGGGCTCCAGGGAGTGTGAGGTATGAACCATTTGTGATAAAAAGAGCTCTTAAAGCT
 GGGAAATAAGTGGTGCTTTATTGAACTCTGGTGACTATCAAGGGAACGCGATGCCCCCTCCCC
 TTCCCTCTCCCTCTCACTTTGGTGGCAAGATCCTTCCTGTCCGTTTGTAGTGATTATAATACT
 GGTCATTTTCTCTCATACATAATCAACCCATTGAAATTTAAATACCACAATCAATGTGAAGCTT
 GAACTCCGTTTAAATATAATACCTATTGTATAAGACCCTTACTGATTCCATTAATGTGCGATTT
 GTTTTAAGATAAACTTCTTTCATAGGTAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTPLLTLSSAERGCPKGCRCEGKMVYCESQKLQEIPSSISAGCLG
LSLRYSLSQKLKYNQFKGLNQLTWLYLDHNNHISNIDENAFNGIRRLKELILSSNRISYFLNNTFR
PVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHLRSNSLRTPVRIHQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKEHLHLEHNQFSKLNALFPRLVSLQNLQWKNISVIGQTMSTWSSLRQLDL
SGNEIEAFSGPSVFQCVPNLQRLNLDNKLTFIQEILDSEWISLNDISLAGNIWECRNICSLVN
WLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVILLVIYVSWKRYPASMKQLQQR
SLMRHRHKKRQSLKQMPSTQEFYVDYKPTNTETSEMLLNGTGPCITYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMQGEAALARLGDGARES
APYRGMVRTALGIIEEGFLKLWQGVTPAIYRHVVYSGGRMVITYEHLREVVFVGKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMMEGKRKLEKPLRFRGVHHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCSGLVASILGTPADVIKSRIMNQP
RDKQGRGLLYKSSTDCLIQAVQGEFMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 127

CGCGGATCGGACCCAAGCAGGTCGGCGGCGGGCGGCAGGAGAGCGGCCGGGCGTCAGCTCCTCGAC
CCCCGTGTCGGGCTAGTCCAGCGAGGCGGACGGGCGGCGTGGGCCCATCGCCAGGCCCGGCATGG
AGCGGTGGCGCGACCGGCTGGCGCTGGTGACGGGGGCGCTCGGGGGGCATCGGCGCGGCCGTGGCC
CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGCGCCCGCACTGTGGGCAACATCGAGGA
GCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGGACTTTGATCCCTACAGATGTGACCTAT
CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATC
TGATCAACAATGCTGGCTTGGCCCGGCTGACACCCTGCTCTCAGGCAGCACCAGTGGTTGGAA
GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGA
AGGAGCGGAATGTGGACGATGGGCACATCATTAACTCAATAGCATGTCTGGCCACCGAGTGTTA
CCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACAGAGGGACT
GAGGCAAGAGCTTCGGGAGGCCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGTGGTGG
AGACACAATTGCGCTTCAAACCTCCACGACAAGGACCCTGAGAAGGCAGCTGCCACCTATGAGCAA
ATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGC'TGTTATCTACGTCCTCAGCACCCCCGCACA
CATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACCTAGTGACTGTGGGAGCTCC
TCCTTCCCTCCCCACCCCTTCATGGCTTGCCCTCCTGCCTCTGGATTTTAGGTGTTGATTTCTGGAT
CACGGGATACCACTTCTGTCCACACCCCGACCAGGGGCTAGAAAATTTGTTTGAGATTTTATA
TCATCTTGTCAAATTGCTTCAGTTGTAAATGTGAAAAATGGGCTGGGGAAAGGAGGTGGTGTCCC
TAATTGTTTTACTTGTTAACTTGTTCTTGTGCCCCCTGGGCACTTGGCCTTTGTCTGCTCTCAGTG
TCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAATCCCCATCTTCTGCACCTCAACGTCGTG
TGGCTCAGGGCTGGGGTGGCAGAGGGAGGCCTTACCTTATATCTGTGTTGTTATCCAGGGCTCC
AGACTTCCTCCTCTGCCTGCCCCACTGCACCCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCC
AGCCCAGTCTTGGCTTCTTGTCCCCTCCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAG
CAGAACACCAGGGCCTGGCCCAGTGGATTTTCATGGTGATCATTAAAAAGAAAAATCGCAACCAA
AAAAAAAAAA

FIGURE 128

MARPGMERWRDRLALVTGASGGIGAAVARALVQGLKVVGCAITVGNIIEELAAECKSAGYPGTLI
PYRCDSLNEEDILSMFSAIRSQHSQVDICINNAGLARPDITLLSGSTSGWKDMFNVNVLALSICTR
EAYQSMKERNVDDGHIININMSGHRVLPVSVTHFYATKYAVTALTEGLRQELREAQTHIRATC
ISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:**Signal peptide:**

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AACTTCTACATGGGCCTCCTGCTGCTGGTGCTCTTCCTCAGCCTCCTGCCGGTGGCCTACACCAT
CATGTCCCTCCCACCTCCTTTGACTGCGGGCCGTTCAAGTGCAGAGTCTCAGTTGCCCGGGAGC
ACCTCCCTCCCGAGGCAGTCTGCTCAGAGGGCCTCGGCCAGAATTCCAGTTCTGGTTTCATGC
CAGCCTGTAAAAGGCCATGGAACCTTTGGGTGAATCACCGATGCCATTTAAGAGGGTTTTCTGCCA
GGATGGAAATGTTAGGTCGTTCTGTGTCTGCGCTGTTCAATTCAGTAGCCACCAGCCACCTGTGG
CCGTTGAGTGCTTGAAATGAGGGAACCTGAGAAAATTAATTTCTCATGTATTTTTCTCATTTATTTA
TTAATTTTTAACTGATAGTTGTACATATTTGGGGTACATGTGATATTTGGATACATGTATACAA
TATATAATGATCAAATCAGGGTAACTGGGATATCCATCACATCAACATTTATTTTTTATTCTTT
TTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCCATCTCAGCTTACTGCAAC
CTCTGCCGTGCCAGGTTCAAGCGATTCTCATGCCTCCACCTCCCAAGTAGCTGGGACTACAGGCAT
GCACCACAATGCCCAACTAATTTTTGTATTTTTAGTAGAGACGGGGTTTTGCCATGTTGCCCAGG
CTGGCCTTGAACTCCTGGCCTCAAACAATCCACTTGCCCTCGGCCTCCCAAAGTGTTATGATTACA
GGCGTGAGCCACCGTGCCCTGGCCTAAACATTTATCTTTCTTTGTGTGGGAACCTTGAAATTAT
ACAATGAATTATTGTTAACTGTCTATCTCCCTGCTGTGCTATGGAACACTGGGACTTCTTCCCTCT
ATCTAACTGTATATTTGTACCAGTTAACCAACCGTACTTCATCCCCACTCCTCTCTATCCTTCCC
AACCTCTGATCACCTCATTCTACTCTCTACCTCCATGAGATCCACTTTTTTAGCTCCCACATGTG
AGTAAGAAAATGCAATATTTGTCTTTCTGTGCCTGGCTTATTTCACTTAACATAATGACTTCCTG
TTCCATCCATGTTGCTGCAAATGACAGGATTTGTTCTTAATTTCAATTAATAAATAACCCACACATG
GCAAAAA

FIGURE 130

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSRGSLLRGPRPRIPVLVSCQPV
KGGHTLGESPMFPKRVFCQDGNVRSFCVCAVHFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTTGTATAAAGACCTCAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
 ATCTTCCTCATCGGGACTAAAATTGGGCTGTTCCCTTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCTG
 TCCATCTGTGTGTCGCTGCGATGCGGGTTTCATTTACTGTAATGATCGCTTTCTGACATCCATTCCAACAG
 GAATACCAGAGGATGCTACAACCTCTACCTTCAGAACAACCAAATAAATAATGCTGGGATTCCTTCAGAT
 TTGAAAACTTGCTGAAAGTAGAAAGAATATACCTATACCACAACAGTTTAGATGAATTCCTACCAACCT
 CCCAAAGTATGTAAGAGAGTTACATTTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTTCAA
 AAATTCCTATCTGGAAGAATTACATTTAGATGACAACCTCTGTCTCTGCAGTTAGCATAGAAGAGGGAGCA
 TTCCGAGACAGCAACTATCTCCGACTGCTTTTCTGTCCCGTAATCACCTTAGCACAAATTCCTGGGGTTT
 GCCCAGGACTATAGAAGAACTACGCTTGGATGATAATCGCATATCCACTATTTTCATCACCATCTCTTCAAG
 GTCTCACTAGTCTAAAACGCCTGGTTCTAGATGGAAACCTGTTGAACAATCATGGTTTAGGTGACAAAGTT
 TTCTTCAACCTAGTTAATTTGACAGAGCTGTCCCTGGTGCAGAAATTCCTGACTGCTGCACCAGTAAACCT
 TCCAGGCACAAACCTGAGGAAGCTTTATCTTCAAGATAACCACATCAATCGGGTGCCCCCAAATGCTTTTT
 CTTATCTAAGGCAGCTCTATCGACTGGATATGTCCAATAATAACCTAAGTAATTTACCTCAGGGTATCTTT
 GATGATTTGGACAATATAACACAACCTGATTCTTCGCAACAATCCCTGGTATTGCGGGTGCAAGATGAAATG
 GGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACGTGCGTGGGCTCATGTGCCAAGCCCCAGAAA
 AGGTTTCGTGGGATGGCTATTAAGGATCTCAATGCAGAACTGTTTGATTGTAAGGACAGTGGGATTGTAAGC
 ACCATTGAGATAACCACTGCAATACCAACACAGTGATCCTGCCAAGGACAGTGGCCAGCTCCAGTGAC
 CAAACAGCCAGATATTAAGAACCCCAAGCTCACTAAGGATCAACAAACCACAGGGAGTCCCTCAAGAAAAA
 CAATTACAATTACTGTGAAGTCTGTACCTCTGATACCATTCATATCTCTTGAAACTTGCTCTACCTATG
 ACTGCTTTGAGACTCAGCTGGCTTAACTGGGCCATAGCCCGGCATTGGGATCTATAACAGAAACAATTGT
 AACAGGGGAACGCAGTGAGTACTTGGTCACAGCCCTGGAGCCTGATTACCCCTATAAAGTATGCATGGTTC
 CCATGGAAACCAGCAACCTCTACCTATTTGATGAAACTCCTGTTTGTATTGAGACTGAAACTGCACCCCTT
 CGAATGTACAACCTTACAACCAACCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAAACCCCAATTTACC
 TTTGGCTGCCATCATTGGTGGGCTGTGGCCCTGGTTACCATTGCCCTTCTTGCTTTAGTGTGTGGTATG
 TTCATAGGAATGGATCGCTCTTCTCAAGGAACCTGTCATATAGCAAAGGGAGGAGAAGAAAGGATGACTAT
 GCAGAAGCTGGCACTAAGAAGGACAACCTCTATCCTGGAAATCAGGGAACCTCTTTTCAGATGTTACCAAT
 AAGCAATGAACCCATCTCGAAGGAGGAGTTGTAATACACACCATATTTCTCCTAATGGAATGAATCTGT
 ACAAAAACAATCACAGTGAAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC
 TCACACTCATGATGCTGAAGGACTCACAGCAGACTTGTGTTTTGGGTTTTTTAAACCTAAGGGAGGTGATG
 GT

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCTTCCCAGAGTCCTTTGCCCAGGCCACCCAGGCTTCTTGCCA
 GCCCTGCCGGGCCACTTGTCTTCATGTCTGCCAGGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG
 TGCAGAGGCAGTCTGGGCTTGGCCAGAGCTCAGGGTGCTGAGCGTGTGACCAGCAGTGAGCAGAG
 GCGGGCCATGGCCAGCCTGGGGCTGCTGCTCCTGCTCTTACTGACAGCACTGCCACCGCTGTGGT
 CCTCCTCACTGCCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCTGATCCTGTCT
 GCGCTGGAGAGAGCCACCGTCTTCTTAGAACAGAGGCTGCCGAAATCAACCTGGATGGCATGGT
 GGGGGTCCGAGTGCTGGAAGAGCAGCTAAAAAGTGTCGGGAGAGTGGGGCCAGGAGCCCTGC
 TGCAGCCGCTGAGCCTGCGCGTGGGGATGCTGGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
 CTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAAGAGAGTTCAGCTGACCCTCCAGCCCGG
 GTTTTGGAAGCTCCACATGCCTGGATCCACACTGATGCCTCCTTGGTGTACCCACGTTCCGGG
 CCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCTGGTGCAGCTGCTGGGAACCGGACG
 GACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCCGGCTGCTC
 AGGCTACTGCCTGTCCACCAACTGCTCTTCTCCTCTGGGCCAGAATGAGGGGATGCACACAGG
 GACCACTCCAACAGAGCCAGGACTATATCAACCTCTTCTGCGCCAACATGATGGACTTGAACCGC
 AGAGCTGAGGCCATCGGATACGCCTACCTTACCCGGGACATCTTCATGGAAAACATCATGTTCTG
 TGAATGGGCGGCTTCTCCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA
 AACAGCAGGAAGGATGCTTCGGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
 TATCAGCAGCATTTTTTCGAGGAGAGTGAAGAGGCGAGAAAAACAATTTCCAGATTCTCGCTCTGT
 TGCTCAGGCTGGAGTACAGTGGCGCAATCTCGGCTCACTGCAACCTTTGCCTCCTGGGTTCAAGC
 AATTCTCTTGCCTCATCCTCCCGAGTAGCTGGGACTACAGGAGCGTGCCACCATACTGGCTAAT
 TTTTATATTTTTTTAGTAGAGACAGGGTTTCATCATGTGCTCATGCTGGTCTCGAACTCCTGAT
 CTCAAGAGATCCGCCCACCTCAGGCTCCCAAAGTGTTGGGATTATTAGGTGTGAGCCACCGTGTCTG
 GCTGAAAAGCACTTTCAAAGAGACTGTGTTGAATAAAGGGCCAAGGTTCTTGCCACCCAGCACTC
 ATGGGGGCTCTCTCCCTAGATGGCTGCTCCTCCACAAACACAGCCACAGCAGTGGCAGCCCTGG
 GTGGCTTCCTATACATCCTGGCAGAATACCCCCAGCAAACAGAGAGCCACACCCATCCACACCG
 CCACCACCAAGCAGCCGCTGAGACGGACGGTTCCATGCCAGCTGCCTGGAGGAGGAACAGACCCC
 TTTAGTCCTCATCCCTTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG
 ATAAGCAAAGCCACCCGACACCCAATCTTGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGG
 GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAAGTTCAACTGCAACTGAAAAA

FIGURE 134

MSARGRWEGGGRRACRGSGLGLARAQGAERTVSSEQRPMASLGLLLLLLLLTALPPLWSSSLPGLD
 TAESKATIADLILSALERATVFLEQRLPEINLDGMVGVRVLEEQLKSVREKWAQEPLLQPLSLRV
 GMLGEKLEAAIQRSYLHLKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE
 RSDVCLVQLLGTGTDSSSEPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGPLQQSQD
 YINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILSWQKQEGCGFG
 EPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFKQFSCILIP
 SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCATGGCGGCCGCCTCTGTGGGGCTTCTTTCCCGTCCTGCTGCTG
CTGCTGCTATCGGGGGATGTCCAGAGCTCGGAGGTGCCCGGGGCTGCTGCTGAGGGATCGGGAGG
GAGTGGGGTCGGCATAGGAGATCGCTTCAAGATTGAGGGGCGTGCAAGTTGTTCCAGGGGTGAAGC
CTCAGGACTGGATCTCGGCGGCCCGAGTGCTGGTAGACGGAGAAGAGCACGTCGGTTTCCTTAAG
ACAGATGGGAGTTTTGTGGTTCATGATATACCTTCTGGATCTTATGTAGTGGAAGTTGTATCTCC
AGCTTACAGATTTGATCCCGTTCGAGTGGATATCACTTCGAAAGGAAAAATGAGAGCAAGATATG
TGAATTACATCAAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCCAAATGAAATCTTCAGGT
CCACCTTCTTACTTTATTAAAAGGGAATCGTGGGGCTGGACAGACTTTCTAATGAACCCAATGGT
TATGATGATGGTTCTTCCTTTATTGATATTTGTGCTTCTGCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGGAAATGGAGCAGTCAATGAATATGCTGAATTCCAACCATGAGTTGCCTGAT
GTTTCTGAGTTCATGACAAGACTCTTCTCTTCAAAATCATCTGGCAAATCTAGCAGCGGCAGCAG
TAAACAGGCAAAAGTGGGGCTGGCAAAAGGAGGTAGTCAGGCCGTCCAGAGCTGGCATTTCAC
AAACACGGCAACACTGGGTGGCATCCAAGTCTTGGAACCGTGTGAAGCAACTACTATAAACTT
GAGTCATCCCGACGTTGATCTCTTACAACGTGTATGTT
AACTTTTTAGCACATGTTTTGTACTTGGTACACGAGAAAACCCAGCTTTCATCTTTTGTCTGTAT
GAGGTCAATATTGATGTCACTGAATTAATTACAGTGTCTATAGAAAATGCCATTAATAAATTAT
ATGAACTACTATACATTATGTATATTAATTAAAACATCTTAATCCAGAAATCAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAA

[illegible]

Important features of the protein:

amino acids 1-23

amino acids 161-182

amino acids 184-187

amino acids 37-40, 236-239

amino acids 151-154

amino acids 33-38, 36-41, 38-44, 229-234

amino acids 238-241

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTTCTGTGAGATTGATTTCTCCCCAGTTCCCCTGTGGGTCTGAGGGGA
 CCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCCAAAACAA
 GTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCTGTTCCAGGC
 CTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACGACCTGTGC
 CACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTCACGGGAGGCTTGGCAGT
 TTTTCTTACTCCTGTGGTCTCCAGATTTCAGGCCTAAGATGAAAGCCTCTAGTCTTGCCCTTCAGC
 CTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAGACACTCAATTTGGG
 AAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCTGAGATACGGGGCAGTG
 TGCAAGCCAAAGATGGAACATTGACATCAGAATCTTAAGGAGGACTGAGTCTTTGCAAGACACA
 AAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTTGCTAAGACTCTATCTGGACAGGGTATTTAA
 AACTACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGCAGCCTCGCCAATTCCTTTCTTA
 CCATCAAGAAGGACCTCCGGCTCTCTCATGCCCACATGACATGCCATTGTGGGGAGGAAGCAATG
 AAGAAATACAGCCAGATTCTGAGTCACTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
 TTTGGGGGAACTAGACATTCTTCTGCAATGGATGGAGGAGACAGAATAGGAGGAAAGTGATGCTG
 CTGCTAAGAATATTCGAGGTCAAGAGCTCCAGTCTTCAATACCTGCAGAGGAGGCATGACCCAA
 ACCACCATCTCTTTACTGTACTAGTCTTGTGCTGGTCACAGTGTATCTTATTTATGCATTACTTG
 CTTCTTGCATGATTGTCTTTATGCATCCCCAATCTTAATTGAGACCATACTTGTATAAGATTTT
 TGTAATATCTTTCTGCTATTGGATATATTTATTAGTTAATATATTTATTTATTTTTTTGCTATTTA
 ATGTATTTATTTTTTTTACTTGGACATGAAACTTTAAAAAAATTCACAGATTATATTTATAACCTG
 ACTAGAGCAGGTGATGTATTTTTATACAGTAAAAAAAACCTTGTAATTCAGAAAGAGTGG
 CTAGGGGGGTATTCAATTGTATTCACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGAT
 ATTTGAAATTGAACCAATGACTACTTAGGATGGGTTGTGGAATAAGTTTTGATGTGGAATTGCAC
 ATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTATCTTCCAG
 CCAGGAATCCTACACGGCCAGCATGTATTTCTACAAATAAAGTTTTCTTTGCATACCAAAAAAA
 AAAAAAAAAA

FIGURE 138

MRQFPKTSFDISPEMSFSIYSLQVPAVPGLTCWALTAEPGWGQNGGATTCATNSHSDSELRPEIF
SSREAWQFFLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
FSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKIS
SLANSFLTICKDLRLSHAHTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGGTCGGTTCCGCATCCAGCC
TAGCGTGTCCACGATCGCGGCTGGGCTCCGGGACTTTCGCTACCTGTTGCGTAGCGATCGAGGTGC
TAGGGATCGCGGTCTTCCCTTCGGGGATTCTTCCCGGCTCCCGTTCGTTCCCTCTGCCAGAGCGGAA
CACGGAGCGGAGCCCCAGCGCCGAACCTCGGCTGGAGCCAGTTCTAACTGGACCACGCTGCC
ACCACCTCTCTTCAGTAAAGTTGTTATTGTTCTGATAGATGCCTTGAGAGATGATTTTGTGTTG
GGTCAAAGGGTGTGAAATTTATGCCCTACACAACCTACCTTGTGGAAAAAGGAGCATCTCACAGT
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAATCAAGGCATTGATGACGGGGAG
CCTTCCTGGCTTTGTGACGTCATCAGGAACCTCAATTCTCCTGCACTGCTGGAAGACAGTGTGA
TAAGACAAGCAAAGCAGCTGGAAGAAAGTATAGTCTTTTATGGAGATGAAACCTGGGTTAAATTA
TTCCCAAAGCATTGTGTGGAATATGATGGAACAACCTCATTTCGTGTGAGATTACACAGAGGT
GGATAATAATGTCACGAGGCATTGGATAAAGTATTAAGAGAGGAGATTGGGACATATTAATCC
TCCACTACCTGGGGCTGGACCACATTGGCCACATTTCAGGGCCCAACAGCCCCCTGATTGGGCAG
AAGCTGAGCGAGATGGACAGCGTGTGATGAAGATCCACACCTCACTGCAGTCGAAGGAGAGAGA
GACGCTTTTACCAATTTGCTGGTTCTTTGTGGTGACCATGGCATGTCTGAAACAGGAAGTCACG
GGGCTCCTCCACCGAGGAGGTGAATACACCTCTGATTTTAATCAGTTCTGCGTTTGAAGGAAA
CCCGGTGATATCCGACATCCAAAGCACGTCCAATAGACGGATGTGGCTGCGACACTGGCGATAGC
ACTTGGCTTACCGATTCCAAAGACAGTGTAGGGAGCCTCCTATTCCCAGTTGTGGAAGGAAGAC
CAATGAGAGAGCAGTTGAGATTTTACATTTGAATACAGTGCAGCTTAGTAAACTGTTGCAAGAG
AATGTGCCGTCATATGAAAAGATCCTGGGTTTGAGCAGTTTAAATGTGAGAAAGATTGCATGG
GAACTGGATCAGACTGTACTTGGAGGAAAAGCATTGAGAAGTCCTATTCAACCTGGGCTCCAAGG
TTCTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTTGTCCCTGAGTGCACAAGTGGCCAG
TTCTCACCTGCTCCTGCTCAGCGTCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAGTCCCA
CTGTCTCTCTGGGTTTTCTCTGCTCTTTATTTGGTGATCCTGGTTCTTTCGGCCGTTACAGT
CATTGTGTGCACCTCAGCTGAAAGTTCGTGCTACTTCTGTGGCCTCTCGTGGCTGGCGGCAGGCT
GCCTTTCGTTTACCAGACTCTGGTTGAACACCTGGTGTGTGCCAAGTGTGGCAGTGCCCTGGAC
AGGGGGCTCAGGGAAGGACGTGGAGCAGCCTTATCCCAGGCTCTGGGTGTCCCGACACAGGTG
TTCACATCTGTGCTGTGAGTGCAGATGCCTCAGTTCTTGAAAGCTAGGTTCTGCGACTGTTAC
CAAGGTGATTGTAAGAGCTGGCGGTACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGTGAA
TCGGACAGCCTCCAGCAGAGGTGTGGGAGCTGCAGCTGAGGGAAGAAGAGACAATCGGCTGGA
CACTCAGGAGGGTCAAAGGAGACTTGGTCGCACCACTCATCTGCCACCCCAAGATGCATCCT
GCCTCATCAGGTCCAGATTTCTTTCCAAGGCGGACGTTTTCTGTGGAATTCTTAGTCTTGGCC
TCGGACACCTTCAATTCGTTAGCTGGGGAGTGGTGGTGAGGCAGTGAAGAAGAGGCGGATGGTCAC
ACTCAGATCCACAGAGCCAGGATCAAGGGACCCACTGCAGTGGCAGCAGGACTGTTGGCCCCC
ACCCCAACCTGCACAGCCCTCATCCCTCTTGGCTTGAGCCGTCAGAGGCCCTGTGCTGAGTGT
CTGACCGAGACACTCACAGCTTTGTCTCAGGGCACAGGCTTCTCGGAGCCAGGATGATCTGTG
CCACGCTTGACCTCGGGCCCATCTGGGCTCATGCTCTCTCCTGCTATTGAATTAGTACCTAG
CTGCACACAGTATGTAGTTACCAAAAGATAAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTTLPPPLF
 SKVVIVLIDALRDDVFVFGSKGVKMPYTTTLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
 VDVIRNLNSPALLEDVIRQAKAAGKRIVFYGDETWVKLFPKHFEYDGTTSFFVSDYTEVDNNV
 TRHLDKVLKRGDWDILILHYLGLDGHIGHISGPN SPLIGQKLEMDSVLMKIHTSLQSKERETPLP
 NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
 275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 142

MLLLLLEYNFPIENNCQHLKTTHTFRVKNLNPKKFSIHDQDHKVLVLD SGNLI AVDPKNYIRPEI
FFALASSLSSASA EKGSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAAQKESARRPFI
FYRAQVGSWNMLESAAHPGWFICTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 33-36

N-myristoylation site.
amino acids 50-55, 87-92

Interleukin-1
amino acids 37-182

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAGGTTCTGACGCGATGAGGAAG
CACCTGAGCTGGTGGTGGCTGGCCACTGTCTGCATGCTGCTCTTCAGCCACCTCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAATCAAGTGAACCGGAAGGCCCTGCCAGCACTGCCAGATCA
CTGAGGCCCAGGTGGCTGAGAACC GCCGGGAGCCTTCATCAAGCAAGGCCGCAAGCTCGACATT
GACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCCGATGGCATCCA
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTTGTCACCGGCTGCATCAATGCCA
CCCAGGCGGCGAACCAGGGGAGTTCCAGAAGCCAGACAACAAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTTGGTTGGAGAGGGGCGCAGGACT
TCGGGTACCATGCACCAGCCAGTGCTCCTCTGCCTTCTGGCTTTGATCTGGCTCATGGTGAAAT
AAGCTTGCCAGGAGGCTGGCAGTACAGAGCGCAGCAGCGAGCAAATCCTGGCAAGTGACCCAGCT
CTTCTCCCCCAAACCCACGCGTGTTCTGAAGGTGCCCAGGAGCGGCGATGCACTCGCACTGCAAA
TGCCGCTCCACGTATGCGCCCTGGTATGTGCCTGCGTTCTGATAGATGGGGGACTGTGGCTTCT
CCGTCACCTCCATTCTCAGCCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAAATGCTTGAT
GAGAAGAACACATCAGGCACTGCGCCACCTGCTTCACAGTACTTCCAACAACCTCTTAGAGGTAG
GTGTATTCCCGTTTTACAGATAAGGAAACTGAGGCCCAGAGAGCTGAAGTACTGCACCCAGCATC
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCTGGCTTGCTAACCCAGGTTTTCTGCTCT
GTCCAATTCCAGAGCTGTCTGGTGATCACTTTATGTCTCACAGGGACCCACATCCAAACATGTAT
CTCTAATGAAATTGTGAAAGCTCCATGTTTAGAAATAAATGAAAACACCTGA

[illegible]

Important features of the protein:

amino acids 1-26

amino acids 157-171

amino acids 98-102, 110-114

amino acids 76-83

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

amino acids 62-66

GCCTTGGCCTCCCAAAGGGCTGGGATTATAGGCGTGACCACCATGTCTGGTCCAGAGTCTCATTTCCTGATGATTTATAGACTCAAAGAAAACTCATGTTCAGAAGCTCTCTTCTCTTCTGGCCTCCTCTCTGTCTTCTTTCCCTCTTTCTTCTTATTTTAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGGAATCTTTTCATTTTGCTTGTCAGTGGGGTAGGTCACTGAGTCTTAGTTTTTATTTTTTGAAATTTCAACTTTCAGATTCAGGGGGTACATGTGAAGGTTGTTTTATGAGTATATTGCATTGATGCTGAGGTTTGGGGT

FIGURE 148

MFRSSLLFWPPLCLLSLFLILISSIYSESKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

GTCTCCGCGTCACAGGAACCTCAGCACCCACAGGGCGGACAGCGCTCCCTCTACCTGGAGACTTGAC
 TCCCGCGCGCCCAACCTGCTTATCCCTTGACCGTCGAGTGTGAGAGATCCTGCAGCCGCCAGTCC
 CGGCCCCCTCTCCGCCCCACCCACCTCCTGGCTCTTCCTGTTTTACTCCTCTTTTCATTGATA
 ACAAAAGCTACAGCTCCAGGAGCCAGCGCCGGCTGTGACCCAAGCCGAGCGTGGAAGAAATGGGGTT
 CCTCGGGACCGGCACTTGGATTCTGGTGTTAGTGCTCCCGATTCAAGCTTTCCCAAACCTGGAGGAA
 GCCAAGACAAATCTCTACATAATAGAGAATTAAGTGCAGAAAGACCTTTGAATGAACAGATTGCTGAA
 GCAGAAGAAGACAAGATTAAAAAACATATCCTCCAGAAAACAAGCCAGGTGAGAGCAACTATTCTTT
 TGTTGATAACTTGAACCTGCTAAAGGCAATAACAGAAAAGGAAAAAATTGAGAAAGAAAGACAATCTA
 TAAGAAGCTCCCCACTTGATAATAAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAACG
 ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAAATTTCAAGATGATCCAGATGGTCTTCA
 TCAACTAGACGGGACTCCTTTAACCGCTGAAGACATTGTCCATAAAATCGCTGCCAGGATTATGAAG
 AAAATGACAGAGCCGTGTTTGACAAGATTGTTTCTAAACTACTTAATCTCGGCCTTATCACAGAAAGC
 CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTTACAAAAATTAATCTCAAAGGAAGCCAACAA
 TTATGAGGAGGATCCCAATAAGCCCACAAGCTGGACTGAGAATCAGGCTGGAAAAATACCAGAGAAAG
 TGACTCCAATGGCAGCAATTCAAGATGGTCTTGCTAAGGGAGAAAACGATGAAACAGTATCTAACACA
 TTAACCTTGACAAATGGCTTGGAAGGAGAACTAAAACCTACAGTGAAGACAACCTTGAGGAACTCCA
 ATATTTCCCAAATTTCTATGCGCTACTGAAAAGTATTGATTGAGAAAAAGAGCAAAAGAGAAAGAAA
 CACTGATTACTATCATGAAAACACTGATTGACTTTGTGAAGATGATGGTGAAATATGGAACAATATCT
 CCAGAAGAAGGTGTTTCTACCTTGAAAACCTGGATGAAATGATTGCTCTTCAGACCAAAAACAAGCT
 AGAAAAAATGCTACTGACAATATAAGCAAGCTTTTCCAGCACCATCAGAGAAGAGTCATGAAGAAA
 CAGACAGTACCAAGGAAGAAGCAGCTAAGATGGAAAAGGAATATGGAAGCTTGAAGGATTCCACAAAA
 GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAAACAGAAGCCTATTTGGAAGC
 CATCAGAAAAAATATGAAATGGTTGAAGAAACATGACAAAAGGGAAATAAAGAAGATTATGACCTTT
 CAAAGATGAGAGACTTCATCAATAAACAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAAGGAA
 GAAGCCGAGGCCATCAAGCGCATTTATAGCAGCCTGTAAAAATGGCAAAAGATCCAGGAGTCTTTCAA
 CTGTTTCAGAAAACATAATATAGCTTAAAACACTTCTAATTCTGTGATTAAAAATTTTTGACCCAAGG
 GTTATTAGAAAGTGCTGAATTTACAGTAGTTAACCTTTTACAAGTGGTTAAACATAGCTTTCTTCCC
 GTAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAAAAAAAAAAAAAAAA

FIGURE 150

MGFLGTGTWILVLVLP IQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG
QSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDPNKPTSWTENQAGKIPEKVTMAAIQDGLAKGENDETVSNTLTLTNGLE
RRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLDIVKMMVKYGTISPEEGV
SYLENLDEMIALQTKNKLEKNATDNISKLFAPSEKSHEETDSTKEEAAKMEKEYGSLKDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKKGNKEDYDLSKMRDFINKQADAYVEKGILDK
EEAEAIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCTGTGGAGCTCAA
CATGGTCTCTGAGTGGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTTATCTGC
ATAATAACCAGCTTCTAGCTGGAGGGCTGCATGCAGGGAAGGTTCATTAAAGGTGAAGAGATCAGC
GTGGTCCCCAATCGGTGGCTGGATGCCAGCCTGTCCCCCGTCATCCTGGGTGTCCAGGGTGAAG
CCAGTGCCTGTGCATGTGGGGTGGGGCAGGAGCCGACTCTAACACTAGAGCCAGTGAACATCATGG
AGCTCTATCTTGGTGCCAAGGAATCCAAGAGCTTCACCTTCTACCGCGGGACATGGGGCTCACC
TCCAGCTTCGAGTCTGGCTGCCTACCCGGGCTGGTTCCCTGTGCACGGTGCCTGAAGCCGATCAGCC
TGTCAGACTCACCCAGCTTCCCGAGAATGGTGGCTGGAATGCCCCCATCACAGACTTCTACTTCC
AGCAGTGTGACTTAGGGCAACGTGCCCCCAGAACTCCCTGGGCAGAGCCAGCTCGGGTGAGGGGT
GAGTGGAGGAGACCCATGGCGGACAATCACTCTCTCTGCTCTCAGGACCCCCACGTCTGACTTAG
TGGGCACCTGACCACTTTGTCTTCTGGTTCCCGAGTTTGATAAATTCTGAGATTTGAGCTCAGT
CCACGGTCTCCCCACTGGATGGTGTCTACTGCTGTGGAACCTTGTA AAAACCATGTGGGGTAA
CTGGGAATAACATGAAAAGATTTCTGTGGGGGTGGGGTGGGGGAGTGGTGGGAATCATTCCTGCT
TAATGGTAAC TGACAAGTGTACCTGAGCCCCGAGGCCAACCCATCCCCAGTTGAGCCTTATA
GGGTGAGTAGCTCTCCACATGAAGTCCTGTCACTCACCCTGTGCAGGAGAGGGAGGTGGTCATA
GAGTCAGGGATCTATGGCCCTTGGCCAGCCCCACCCCTTCCCTTTAATCCTGCCACTGTCTATA
TGCTACCTTTCTATCTCTTCCCTCATCATCTTGTGTGGGCATGAGGAGGTGGTGTGTGAGAA
GAAATGGTCTGAGCTCAGAAGATAAAAGATAAGTAGGGTATGCTGATCCTCTTTTAAAAACCCAA
GATACAATCAAAATCCCAGATGCTGGTCTCTATTCCCATGAAAAAGTGTCTGACATATTGAGA
AGACCTACTTACAAAGTGGCATATATTGCAATTTATTTTAAATTAAGATACCTATTTATATATT
TCTTTATAGAAAAAGTCTGGAAGAGTTTACTTCAATTGTAGCAATGTGAGGGTGGTGGCAGTAT
AGGTGATTTTTCTTTTAAATCTGTGTAATTTATCTGTATTTCTTAATTTTCTACAATGAAGATGA
ATTCTTGTATAAAAAATAAGAAAAGAAATTAATCTTGAGGTAAGCAGAGCAGACATCATCTCTGA
TTGTCTCAGCCTCCACTTCCCCAGAGTAAATTCAAATTGAATCGAGCTCTGCTGCTCTGGTTGG
TTGTAGTAGTGATCAGGAACAGATCTCAGCAAAGCCACTGAGGAGGAGGCTGTGCTGAGTTTGT
GTGGCTGGAATCTCTGGGTAAAGAACTTAAAGAACAAAAATCATCTGGTAATCTTTCTCTAGAAG
GATCACAGCCCCCTGGGATTTCAAGGCATTGGATCCAGTCTCTAAGAAGGCTGCTGTACTGGTTGA
ATTGTGTCCCCCTCAAATTCACATCCTTCTTGGAATCTCAGTCTGTGAGTTTATTTGGAGATAAG
GTCTCTGCAGATGTAGTTAGTTAAGACAAGGTCTGCTGGATGAAGGTAGACCTAAATTCATAT
GACTGGTTTTCTGTATGAAAAGGAGAGGACACAGAGACAGAGGAGACGCGGGGAAGACTATGTA
AAGATGAAGGCAGAGATCGGAGTTTTGCGCCACAAGCTAAGAAACACCAAGGATTGTGGCAACC
ATCAGAAGCTTGGAAGAGGCAAAGAAGAAATCTTCCCTAGAGGCTTTAGAGGATAACGGCTCTG
CTGAAACCTTAATCTCAGACTTCCAGCCTCCTGAACGAAGAAAGATAAATTTCTGGCTGTTTTAA
GCCACCAAGGATAATTGGTTACAGCAGCTCTAGGAACTAATACAGCTGCTAAAATGATCCCTGT
CTCCTCGTGTTTACATTCTGTGTGTGTCCCTCCCACAATGTACCAAAGTTGTCTTTGTGACCAA
TAGAATATGGCAGAAGTGATGGCATGCCACTTCCAAGATTAGGTTATAAAAGACACTGCAGCTTC
TACTTGAGCCCTCTCTCTCTGCCACCCACCGCCCCCAATCTATCTTGGCTCACTCGCTCTGGGGG
AAGCTAGCTGCCATGCTATGAGCAGGCCATATAAGAGACTTACGTGGTAAAAAATGAAGTCTCCT
GCCACAGCCACATTAGTGAACTTAGAAGCAGAGACTCTGTGAGATAATCGATGTTTGTGTTTT
AAGTTGCTCAGTTTTGGTCTAACTTGTATGCAGCAATAGATAAATAATATGCAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEESISVVPNRWLDASLSPVILGVQGG
 QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP
 VRLTQLPENGGWNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCAATGCGCCGC
 CCTGCAGAAATCTGTGAGCTCTTTCCCTTATGGGGACCCTGGCCACCAGCTGCCTCCTTCTCTTGG
 CCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCCAGCTGCAGGCTTGACAAGTCCAAC
 TTCCAGCAGCCCTATATCACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA
 CAACACAGACGTTTCGTCTCATTGGGGAGAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
 ATCTGATGAAGCAGGTGCTGAACCTTCACCCTTGAAGAAGTGCTGTTCCCTCAATCTGATAGGTTT
 CAGCCTTATATGCAGGAGGTGGTGGCCTTCTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCA
 TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAAGC
 TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAAGTGGATTTGCTGTTTATGTCTCTGAGAAAT
 GCCTGCATTGACACAGAGCAAAGCTGAAAAATGAATAACTAACCCCTTTCCCTGCTAGAAATAA
 CAATTAGATGCCCCAAAGCGATTTTTTTTAAACAAAAGGAAGATGGGAAGCCAACTCCATCATG
 ATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAAACCAATGCCACTTTTGTTTATA
 AGACCAGAAGGTAGACTTTCTAAGCATAGATATTTATTGATAACATTTTCATTGTAAGTGGTGTTC
 TATACACAGAAAAACAATTTATTTTTTAAATAATTGCTTTTTTCCATAAAAAAGATTACTTTCCAT
 TCCTTTAGGGGAAAAAACCCCTAAATAGCTTCATGTTTCCATAATCAGTACTTTATATTTATAAA
 TGTATTTATTATTATTATAAGACTGCATTTTATTTATATCATTTTATTAATATGGATTTATTTAT
 AGAAACATCATTCGATATTGCTACTTGAGTGTAAGGCTAATATTGATATTTATGACAATAATTAT
 AGAGCTATAACATGTTTATTTGACCTCAATAAACACTTGGATATCCC

FIGURE 154

MAALQKSVSSFLMGTLATSCLLLLALLVQGGAAPISSHCRLDKSNFQQPYITNRTFMLAKEASL
ADNNTDVRLIGEKLFGVSMSERCYLMKQVLNFTLEEVLFPQSDRFQPYMQEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:**Signal peptide:**

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCTGT
 CAGTCAGTGCCCGACTTGTGACTGAGTGTGTCAGTGCCCGAGCATGTACCAGGTTCAGTCAGAGGGC
 TGCTTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGC
 TGCCAGGTTTGGGGCTGGGGGCCAAGTGGAGTGAGAACTGGGATCCCAGGGGGAGGGTGCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTCATTAGCCTTTTCTACAGGTGGTTGCAT
 TCTTGGAATGGTCATGGGAACCCACACCTACAGCCACTGGCCCAGCTGCTGCCCCAGCAAAGGG
 CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCTCCCCTAGAGCCTGCTAG
 GCCCAACCGCCACCCAGAGTCCTGTAGGGCCAGTGAAGATGGACCCCTCAACAGCAGGGCCATCT
 CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCCAGGACCTGTACCACGCCCCGT
 TGCTGTGCCCCGCACTGCGTCAGCCTACAGACAGGCTCCACATGGACCCCGGGGCAACTCGGA
 GCTGCTCTACCACAACCAGACTGTCTTCTACAGGCGGCCATGCCATGGCGAGAAGGGCACCCACA
 AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTTCCTTAGCTTGTGTGTGTGTGCGGCCCCGT
 GTGATGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTTTTGGGAAACCTGGAGCCAGGTGTACA
 ACCACTTGCCATGAAGGGCCAGGATGCCCAGATGCTTGGCCCCGTGTGAAGTGCTGTCTGGAGCAG
 CAGGATCCCGGGACAGGATGGGGGGCTTTGGGAAAACCTGCACTTCTGCACATTTTGAAAAGAG
 CAGCTGCTGCTTAGGGCCGCCGGAAGCTGGTGTCTCTCATTTTCTCTCAGGAAAGGTTTTCAA
 GTTCTGCCCATTTCTGGAGGCCACCACTCCTGTCTCTTCTCTTTTCCCATCCCCTGCTACCTG
 GCCCAGCACAGGCACTTTCTAGATATTTCCCCCTTGCTGGAGAAGAAAGAGCCCCCTGGTTTTATT
 TGTTTGTTTACTCATCACTCAGTGAGCATCTACTTTGGGTGCATTCTAGTGTAGTTACTAGTCTT
 TTGACATGGATGATTCTGAGGAGGAAGCTGTTATTGAATGTATAGAGATTTATCCAAATAAATAT
 CTTTATTTAAAAATGAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGHTYSHWPSCCPSKGQDTSEELLRWSTVPVPPLEPA
 RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
 ELLYHNQTVFYRRPCHGEKGTHKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCGATGTCGCTCGTGCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAGCC
GACCGTTCAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACAGTGTGCAACAGGGGACTATTCA
ATTTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCGCTTGTTGAAGGCCACCAA
GATTTGTGTGACGGGCAAAAGCAACTTCCAGTCCTACAGCTGTGTGAGGTGCAATTACACAGAGG
CCTTCCAGACTCAGACCAGACCCTCTGGTGGTAAATGGACATTTTCCTACATCGGCTTCCCTGTA
GAGCTGAACACAGTCTATTTTCATTGGGGCCATAATATTCCTAATGCAAATATGAATGAAGATGG
CCCTTCCATGTCTGTGAATTTACCTCACCAGGCTGCCTAGACCACATAATGAAATATAAAAAA
AGTGTGTCAAGGCCGGAAGCCTGTGGGATCCGAACATCACTGCTTGTAAGAAGAATGAGGAGACA
GTAGAAGTGAACCTTACAACCACTCCCCTGGGAAACAGATACATGGCTCTTATCCAACACAGCAC
TATCATCGGGTTTTCTCAGGTGTTTGAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGA
TTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTTTCCTACTTGTGGC
AGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCCCACAAACAGGCGTCCCTTTCCCTCT
GGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTCCTGCTGTCTCTGCTGGTGGCCA
CATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTTCCTTT
TCTACCACCACACTACTGCCCCCATTAAGGTTCTTGTGGTTTACCCATCTGAAATATGTTTCCA
TCACACAATTTGTTACTTCACTGAATTTCTTCAAACCATTCAGAAAGTGAGGTCACTCCTTGAAA
AGTGGCAGAAAAAGAAAATAGCAGAGATGGGTCCAGTGCAGTGGCTTGCCACTCAAAAGAAGGCA
GCAGACAAAGTCGTCTTCCTTCTTTCCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAA
GAGCGAGGGCAGTCCCAGTGAGAACTCTCAAGACCTCTTCCCCCTTGCCTTAACCTTTTCTGCA
GTGATCTAAGAAGCCAGATTCATCTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACA
AAAGACGATTACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTT
CTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTGAGCAGGAAAAAGATCACAAGCCTGCCACG
ATGGCTGCTGCTCCTTGTAG

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLPGLDLRLRVEPVTTSVATGDYSILMNVS
 LRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIP
 NANMNEDGPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKKNEETVEVNFTTTPLGNRYMALIQH
 STIIGFSQVFEPHQKKQTRASVVIPTVTDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPTGVPFPLDNNK
 SKPGGWLPLLLLSLLVATWVLVAGIYLMWRHERIKKTSFSTTTLLPPIKVLVVPSEICFHHTICYFTEFL
 QNHCRSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLLSNDVNSVCDGTCGKSEGSPSENSQDLFPLA
 ENLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVC PKYHLMKDATAFCAELLHVKKQVSAGKRSQACHD
 GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
 - 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

FIGURE 159

AGCCACCAGCGCAACATGACAGTGAAGACCCTGCATGGCCCAGCCATGGTCAAGTACTTGCTGCT
 GTCGATATTGGGGCTTGCCTTTCTGAGTGAGGCGGCAGCTCGGAAAATCCCCAAAGTAGGACATA
 CTTTTTCCAAAAGCCTGAGAGTTGCCC GCCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
 ATCATCAATGAAAACCAGCGCGTTTCCATGTCACGTAACATCGAGAGCCGCTCCACCTCCCCCTG
 GAATTACACTGTCACTTGGGACCCCAACCGGTACCCCTCGGAAGTTGTACAGGCCCAGTGTAGGA
 ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTTCCCATCCAGCAA
 GAGACCCTGGTCGTCCGAGGAAGCACCAAGGCTGCTCTGTTTCTTCCAGTTGGAGAAGGTGCT
 GGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCATCCACCATGTGCAGTAAGAGGTGCATATCC
 ACTCAGCTGAAGAAG

FIGURE 160

MTVKTLHGPMVKYLLLSILGLAFLSEAAARKIPKVGHTFFQKPESCPPVPGGSMKLDIGIINEN
QRVMSRNIERSRSTSPWNYTVTWDPNRYPSEVVQAQCRNLGCINAQKEDISMNSVPIQQETLVV
RRKHQGCVSFQLEKVLVTVGCTCVTPVIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

ACACCTGGCCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGAGTCAGGACTCCCAGG
 ACAGAGAGTGCACAACTACCCAGCACAGCCCCCTCCGCCCCCTTGGAGGCTGAAGAGGGATTTC
 CAGCCCCTGCCACCCACAGACACGGGCTGACTGGGGTGTCTGCCCCCTTGGGGGGGGGAGCAGC
 AGGGCCTCAGGCCTGGGTGCCACCTGGCACCTAGAAGATGCTGTGCCCTGGTTCTTGCTGTCTCT
 TGGCACTGGGGCCGAAGCCAGTGGTCTTTCTCTGGAGAGGCTTGTGGGGCCTCAGGACGCTACC
 CACTGCTCTCCGGGCTCTCCTGCCGCTCTGGGACAGTGACATACTCTGCCTGCTGGGGACAT
 CGTGCTGCTCCGGGCCCCGTGCTGGCGCCTACGCACCTGCAGACAGAGCTGGTGCTGAGGTGCC
 AGAAGGAGACCGACTGTGACCTCTGTCTGCGTGTGGCTGTCCACTTGGCCGTGCATGGGCACTGG
 GAAGAGCCTGAAGATGAGGAAAAGTTTGGAGGAGCAGCTGACTCAGGGGTGGAGGAGCCTAGGAA
 TGCCTCTCTCCAGGCCAAGTCGTGCTCTCCTTCCAGGCCTACCCTACTGCCCGCTGCGTCTCTGC
 TGGAGGTGCAAGTGCCTGCTGCCCTTGTGCAGTTTGGTCAGTCTGTGGGCTCTGTGGTATATGAC
 TGCTTCGAGGCTGCCCTAGGGAGTGAGGTACGAATCTGGTCTTACTCAGCCCAGGTACGAGAA
 GGAACCTCAACACACACAGCAGCTGCCCTGCCCTGGCTCAACGTGTGACGAGATGGTGACA
 ACGTGATCTGGTTCTGAATGTCTCTGAGGAGCAGCACTTCGGCCTCTCCCTGTACTGGAATCAG
 GTCCAGGGCCCCCAAAACCCCGGTGGCACAAAAACCTGACTGGACCGCAGATCATACCTTGAA
 CCACACAGACCTGGTTCCCTGCCTCTGTATTGAGGTGTGGCCTCTGGAACCTGACTCCGTTAGGA
 CGAACATCTGCCCTTTCAGGGAGGACCCCGCGCACACCAGAACCTCTGGCAAGCCGCCGACTG
 CGACTGCTGACCTGCAGAGCTGGCTGCTGGACGCACCGTGTGCTGCTGCCCGCAGAAGCGGCACT
 GTGCTGGCGGGCTCCGGGTGGGGACCCCTGCCAGCCACTGGTCCCACCGCTTTCTGGGAGAACG
 TCACTGTGGACAAGTTCTCGAGTTCCCATGCTGAAAGGCCACCCTAACCTCTGTGTTGAGGTG
 AACAGCTCGGAGAAGCTGCAGCTGCAGGAGTGCTTGTGGGCTGACTCCCTGGGGCTCTCAAAGA
 CGATGTGCTACTGTTGGAGACACGAGGCCCCAGGACAACAGATCCCTCTGTGCCCTTGAACCCA
 GTGGCTGTACTTCACTACCCAGCAAAGCCTCCACGAGGGCAGCTCGCCTTGGAGAGTACTTACTA
 CAAGACCTGCAGTCAGGCCAGTGTCTGCAGCTATGGGACGATGACTTGGGAGCGCTATGGGCCTG
 CCCCATGGACAAAATACATCCACAAGCGCTGGGCCCTCGTGTGGCTGGCCTGCCTACTCTTTGCCG
 CTGCGCTTTCCCTCATCCTCCTTCTCAAAAAGGATCACGCGAAAGGGTGGCTGAGGCTCTTGAAA
 CAGGACGTCCGCTCGGGGGCGGCCAGGGGCCGCGGGCTCTGCTCCTCTACTCAGCCGATGA
 CTCGGGTTTTGAGCGCCTGGTGGGCGCCCTGGCGTCCGGCCTGTGCCAGCTGCCGCTGCGCGTGG
 CCGTAGACCTGTGGAGCCGTGCTGAACAGAGCGCGCAGGGGCCCGTGGCTTGGTTTACGCGCAG
 CGGCGCCAGACCTGCAGGAGGGCGGCGTGGTGGTCTTGTCTTCTCTCCCGGTGCGGTGGCGCT
 GTGCAGCGAGTGGCTACAGGATGGGGTGTCCGGGCCCGGGGCGCACGGCCCGCACGACGCTTCC
 GCGCCTCGCTCAGCTGCGTGCTGCCCGACTTCTTGCAGGGCCGGGCGCCCGGCGAGCTACGTGGGG
 GCCTGCTTCGACAGGCTGTCCACCCGACGCGGTACCCGCCCTTTCCGACCGTGGCCGCTCTT
 CACACTGCCCTCCCACTGCCAGACTTCTTGGGGGCCCTGCAGCAGCCTCGCGCCCCGCTTCCG
 GCGGCTCCAAGAGAGAGCGGAGCAAGTGTCCCGGGCCCTTACGCCAGCCCTGGATAGCTACTTC
 CATCCCCGGGGACTCCCGCGCCGGGACGCGGGTGGGACCAGGGGCGGGACCTGGGGCGGGGA
 CGGGACTTAAATAAAGGCAGACGCTGTTTTCTAAAAAA

FIGURE 162

MPVPWFLLSLALGRSPVVLSELRLVGPDATHCSPGLSCLWSDILCLPGDIVPAPGPFVLAPTHQLQTELV
 LRCQKETDCDLCLRVAVHLAVHGHWEPEDEEKFGGAADSGVEEPRNASLQAQVVLSTFQAYPTARCVLLEV
 QVPAALVQFGQSVGSVVYDCFEAALGSEVRIWSYTPRYEKEINHTQQLPALPWLNVSAADGNVHLVLNV
 EEQHFGLSLYWNQVQGPPKPRWHKNLTGPQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQN
 LWQAARLRLTLQSWLLDAPCSLPAAEALCWRAPGGDPCQPLVPPLSWENVTVDKVLEFPLKKGHPNLCVQ
 VNSSEKLQLQECLWADSLGPLKDDVLLLETRGPDNRSLEPSGCTSLPSKASTRAARLGEYLLQDLQS
 GQCLQLWDDDLGALWACPMCKYIHKRWALVWLACLLFAAALSLILLKKDHAKGWLRLKQDVRSAAAARG
 RAALLYSADDSGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAVFHAQRRQTLQEGGVVLLFSP
 GAVALCSEWLQDGVSGPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLLHPDAVPALFRTVPVFT
 LPSQLPDFLQPRAPRSGRLQERAEQVSRAALQPALDSYFHPPGTPAPGRGVGPGAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,
 334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,
 692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTTGACTGTGGGATCCCTGGCT
 GCTCACGCCCCCTGAGGACCCCTCGGATCTGCTCCAGCACGTGAAATTCAGTCCAGCAACTTTGA
 AAACATCCTGACGTGGGACAGCGGGCCAGAGGGCACCCAGACACGGTCTACAGCATCGAGTATA
 AGACGTACGGAGAGAGGGACTGGGTGGCAAAGAAGGGCTGTCAGCGGATCACCCGGAAGTCTGTC
 AACCTGACGGTGGAGACGGGCAACCTCACGGAGCTCTACTATGCCAGGGTCACCGT
 GTCAGTGGGGAGGCCGGTCAGCCACCAAGATGACTGACAGGTTGAGTCTCTGACGACACTAC
 CCTAAGCCACCTGATGTGACCTGTATCTCCAAAGTGAGATCGATTGAGATGATTGTTTCATCCTA
 CCCCCACGCCAATCCGTGACGGCGATGGCCACCGGCTAACCTGGAGACATCTTCCATGACCTG
 TTCTACCACTTAGAGCTCCAGGTCAACCGCACCTACCAAATGCACCTTGGAGGGAAGCAGAGAGA
 ATATGAGTCTTCGGCCTGACCCCTGACACAGAGTTCCTTGGCACCATCATGATTGCGTTCCCA
 CCTGGGCCAAGGAGAGTGGCCCTACATGTGCCGAGTGAAGACACTGCCAGACCGGACATGGACC
 TACTCCTTCTCCGGAGCCTTCTGTCTCCATGGGCTTCTCGTCGAGTACTCTGCTACCTGAG
 CTACAGATATGTACCAAGCCGCTGCACCTCCCAACTCCCTGAACGTCCAGCGAGTCTGACTT
 TCCAGCCGCTGCGCTTTCATCCAGGAGCACGTCTGATCCCTGTCTTTGACCTCAGCGGCCACG
 AGTCTGGCCACGCTGTCCAGTACTCCAGATCAGGGTGTCTGGACCCAGGGAGCCCGCAGGAGC
 TCCACAGCGGCATAGCCTGTCCGAGATCACCTACTTAGGGCAGCCAGACATCTCCATCCTCCAGC
 CCTCCAACGTGCCACCTCCCGAGATCCTTCCCCACTGTCTATGCCCAAACGCTGCCCTTGAG
 GTCGGGGCCCCATCCTATGCACCTCAGGTGACCCCGAAGCTCAATTCCCATTTACGCCCCACA
 GGCCATCTTAAGGTCCAGCCTTCTCCTATGCCCTCAAGCCACTCCGGACAGCTGGCCTCCCT
 CCTATGGGGTATGTCATGGAAGGTTCTGGCAAAGACTCCCCCACTGGGACACTTTCTAGTCTTAA
 CACCTTAGGCCTAAAGGTGAGCTTTCAGAAAGAGCCACAGCTGGAAGCTGCATGTTAGGTGGCCT
 TTCTCTGCAGGAGGTGACCTCCTTGGCTATGGAGGAATCCCAAGAAGCAAAATCATGACCAGC
 CCCTGGGGATTGTCACAGACAGAATCTGACCCAAATGTGCTACACAGTGGGGAGGAAGGACAC
 CCACAGTACCTAAAGGGCCAGCTCCCCCTCCTCTCCTCAGTCCAGATCGAGGGCCACCCCATGTC
 CCTCCCTTTGCAACCTCCTTCCGGTCCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGGCC
 TGCTGGAGTCCCTTGTGTGTCCCAAGGATGAAGCCAAGAGCCAGCCCTGAGACCTCAGACCTG
 GAGCAGCCACAGAATGGATTCTCTTTTCAGAGGCCTGGCCCTGACTGTGAGTGGGAGTCTTG
 AGGGGAATGGGAAAGGCTTGGTGCTTCCCTGTCCTTACCCAGTGTACATCCTTGGCTGTCA
 ATCCCATGCCCTGCCCATGCCACACACTCTGCGATCTGGCCTCAGACGGGTGCCCTTGAGAGAAGC
 AGAGGGAGTGGCATGCAGGGCCCTGCCATGGGTGCGCTCCTCACCGGAACAAAGCAGCATGATA
 AGGACTGCAGCGGGGAGCTCTGGGGAGCAGCTTGTGTAGACAAGCGGTGCTCGCTGAGCCCTG
 CAAGGCAGAAATGACAGTGCAGGAGGAAATGCAGGGAAACTCCCGAGTCCAGAGCCCCACCTC
 CTAACACCATGGATTCAAAGTGCTCAGGGAATTTGCCTCTCCTTGCCCCATTCTTGCCAGTTTC
 ACAATCTAGCTCGACAGAGCATGAGGCCCTGCCTCTTCTGTATTGTTCAAAGGTGGGAAGAGA
 GCCTGGAAAAGAACAGGCCCTGGAAAAGAACAGAAGGAGGCTGGGCAGAACAGAACACCTGC
 ACTTCTGCCAAGGCCAGGGCCAGCAGGACGGCAGGACTCTAGGGAGGGGTGTGGCCTGCAGTCA
 TTCCAGCCAGGGCAACTGCCTGACGTTGCACGATTTGAGCTTCATTCTCTGATAGAACAAAGC
 GAAATGCAGGTCCACCAGGGAGGGAGACACACAAGCCTTTTCTGCAGGCAGGAGTTTCAGACCT
 ATCCTGAGAATGGGTTTGAAGGAAGGTGAGGGCTGTGGCCCTGGACGGGTACAATAACACAC
 TGTAAGTGTGTCACAACTTTGCAAGCTCTGCCTTGGGTTCAAGCCATCTGGGCTCAAATTCAGC
 CTCACCACTCACAGCTGTGTGACTTCAAACAAATGAAATCAGTGCCAGAACCTCGGTTCCCTC
 ATCTGTAATGTGGGGATCATAACACCTACCTCATGGAGTTGTGGTGAAGATGAAATGAAGTCATG
 TCTTTAAAGTGCTTAATAGTGCTGGTACATGGGCAGTGCCCAATAACGGTAGCTATTTAAAAA
 AAAAAAA

FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTPVYSIEYKTYGERDW
 VAKKGCQRITRKSCNLTVETGNLTELYARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVTCIS
 KVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHGGKQREYEFFGLTPDT
 EFLGTIMICVPTWAKESAPYMCRVKTLDPRTWTYSFSGAFLFSMGFLVAVLCYLSYRYVTKPPAP
 PNSLNVQRVLTQPLRFIQEHVLI PVFDLSGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT
 YLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFPFYAPQAISKVQPSSY
 APQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPGQLQKEPPAGSCMLGGLSLQEVTSIAM
 EESQEAQSLHQPLGICTDRTPDNVLHSGEETPQYLGKQLPLLSSVQIEGHPMSLPLQPPSGPC
 SPSPQGPSPWGLLESVCPKDEAKSPAPETSDLEQPTELDLSLFRGLALTQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAAAAAAGTCACCCGGGCCCCGGGTGGCCACAACATGG
CTGCGGCGCCGGGGCTGCTCTTCTGGCTGTTTCGTGCTGGGGGCGCTCTGGTGGGTCCC GGCCAG
TCGGATCTCAGCCACGACGGCGTTTCTCGGACCTCAAAGTGTGCGGGGACGAAGAGTGCAGCAT
GTTAATGTACCGTGGGAAAGCTCTTGAAGACTTCACGGGCCCTGATTGTCGTTTTGTGAATTTTA
AAAAAGGTGACGATGTATATGTCTACTACAACTGGCAGGGGGATCCCTTGAAC TTTGGGCTGGA
AGTGTTGAACACAGTTTTTGGATATTTTCCAAAAGATTTGATCAAGGTACTTCATAAATACACGGA
AGAAGAGCTACATATTCAGCAGATGAGACAGACTTTGTCTGCTTTGAAGGAGGAAGAGATGATT
TTAATAGTTATAATGTAGAAGAGCTTTTAGGATCTTTGGAAGTGGAGGACTCTGTACCTGAAGAG
TCGAAGAAAGCTGAAGAAGTTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTCGGGGGCGTGA
ACTTGACCCTGTGCCTGAGCCCCGAGGCATTTCAGAGCTGATTTCAGAGGATGGAGAAGGTGCTTTCT
CAGAGAGCACCGAGGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGCCACCCTCACACCAGCGGT
CCTGCGGCTAACGCTCAGGGAGTGCAGTCTTCGTTGGACACTTTTGAAGAAATTCGCACGATAA
ATTGAAAGTGCCGGAAGCGAAAGCAGAACTGGCAATAGTTCTCCTGCCTCGGTGGAGCGGGAGA
AGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGC GTTATT
CATTACAGCAAAGGATTTTCGTTGGCATCAAAATCTAAGTTTGT TTTTACAAAGATTGTTTTTAGTA
CTAAGCTGCCTTGGCAGTTTGCATTTTTTGAGCCAAACAAAATATATTATTTTCCCTTCTAAGTA
AAAAAAAAAAAAAAAAAAAA

FIGURE 166

MAAAPGLLFWLFLVGLALWWVPGQSDLGRRFSDLKVCGDEECMLMYRGKALEDTGPDCRFVN
 FKKGDDVYVYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEELHIPADETDFVCFEGGRD
 DFNSYNVEELLGSLELEDSVPEESKKAEEVSQHREKSPEESRGRELDPVPEPEAFRAEDGEDGA
 FSESTGLQGQPSAQESHPTSGPAANAQGVQSSLDTFEEILHDKLKVPGSESRTGNSSPASVER
 EKTDAYKVLKTEMSQRGSGQCVIHYSKGFRWHQNLSEFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTCTCAGAGGCCGGGGAAGAGAAGCAAAGCGC
 AACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCCCTCTAGGACATACACGGGACCCCTAACTTC
 AGTCCCCCAAACGCGCACCTCGAAGTCTTGAAGTCCAGCCCCGCACATCCACGCGCGGCACAGG
 CGCGGCAGGCGGCAGGTCCCGGCCGAAGGCGATGCGCGCAGGGGGTCGGGCAGCTGGGCTCGGGC
 GGCGGGAGTAGGGCCCCGGCAGGGAGGCAGGGAGGCTGCATATTCAGAGTCCGCGGGCTGCGCCCTG
 GGCAGAGGCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGCGATGAGCCGCGTGG
 TCTCGCTGCTGCTGGGCGCCGCGCTGCTGCTGCCGCCACGGAGCCTTCTGCCGCCGCGTGGTCAAGC
 GGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTACAAAATGGCCTACTTCCATGAAGT
 GTCCAGCCGAGTGAGCTTTCAGGAGGCACGCCCTGGCTTGTGAGAGTGAGGGAGGAGTCCCTCCTCA
 GCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAAACCTGACAAAACCCGGG
 ACAGGGATTCTGATGGTGATTTCTGGATAGGGCTTTGGAGGAATGGAGATGGGCAACATCTGG
 TGCCTGCCCAGATCTCTACCAGTGGTCTGATGGAAGCAATTCAGTACCGAACTGGTACACAG
 ATGAACCTTCTGCGGAAGTGAAAAGTGTGTTGTGATGTATCACCACCAACTGCCAATCTTGGC
 CTTGGGGGTCCCTACCTTTACCAGTGGATGATGACAGGTGTAACATGAAGCACAAATATATTTG
 CAAGTATGAACCAGAGATTAATCCAACAGCCCCGTGAGAAAAGCCTTATCTTACAAATCAACCAG
 GAGACACCCATCAGAATGTGGTTGTTACTGAAGCAGGTATAATTCCTCAATCTAATTTATGTTGTT
 ATACCAACAATACCCCTGCTCTTACTGATACTGGTTGCTTTTGGAACCTGTTGTTTCCAGATGCT
 GCATAAAAGTAAAGGAAGAACAACAACTAGTCCAAACAGTCTACACTGTGGATTTCAAAGAGTA
 CCAGAAAAGAAAGTGGCATGGAAGTATAATAACTCATTGACTTGGTTCCAGAATTTTGTAATCT
 GGATCTGTATAAGGAATGGCATCAGAACATAGCTTGAATGGCTTGAAATCACAAAGGATCTGC
 AAGATGAAGTGAAGCTCCCCCTTGAGGCAATATTAAGTAATTTTATATGTCTATTATTTCA
 TTTAAAGAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTTTGCTAAAGGATGCACCCAA
 ACTTCAAACCTTCAAGCAATGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTCGGGAGTA
 TGTGTGTTAGAAGCAATTCCTTTTATTTCTTTCACCTTTCATAAGTTGTTATCTAGTCAATGTAA
 TGTATATTGTATTGAAATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAAGTGTGATAAAA
 ATGAAGTGTCTAATATTTATTTTATGGCATCTCATTTTCAATACATGCTCTTTTGATTAAAG
 AAACCTTATTACTGTTGTCAACTGAATTCACACACACACAAATATAGTACCATAGAAAAAGTTTGT
 TTTCTCGAAATAATTCATCTTTTCTGCTTTTGGTCAATGTCTAGGAAATCTCTTCAGA
 AATAAGAAGCTATTTTCATTAAGTGTGATATAAACCTCCTCAAACATTTTACTTAGAGGCAAGGAT
 TGTCTAATTTCAATTTGTGCAAGACATGTGCCCTTATAATTTTATAGCTTAAATTAACAGATT
 TTGTAATAATGTAACCTTTGTTAATAGGTGCATAAACACTAATGCAGTCAATTTGAACAAAAGAAG
 TGACATACACAATATAAATCATATGTCTTCACACGTTGCCATATAATGAGAAGCAGCTCTCTGA
 GGGTTCTGAAATCAATGTGGTCCCTCTCTTGCCCACTAAACAAAGATGGTTGTTCCGGGGTTGGG
 ATTGACACTGGAGGCAGATAGTTGCAAGTTAGTCTAAGGTTTCCCTAGCTGTATTAGCCTCTG
 ACTATATTAGTATACAAAGAGGTGATGTGGTTGAGACCAGGTGAATAGTCACTATCAGTGTGGAG
 ACAAGCACAGCACACAGACATTTTAGGAAGGAAAGGAAGTACGAAATCGTGTGAAAATGGGTTGG
 AACCATCAGTGATCGCATATTCATTGATGAGGGTTTGTCTGAGATAGAAAATGGTGGCTCCTTT
 CTGTCTTATCTCCTAGTTTCTTCAATGCTTACGCCCTTGTCTTCTCAAGAGAAAGTTGTAAGTCT
 CTGGTCTTCATATGTCCCTGTGCTCCTTTTAACCAATAAAGAGTTCTTGTCTTCTGGGGGAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVVSLLLGAALLCGHGAFCRRVVSQKVCFADFKHPCYKMAYFHELSSRVSFQEARLACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDGFWIGLWRNGDGQTSGACPDLYQWSDGSNSQ
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKS KGRTKTSPNQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-
145, 212-217